

د. معتوق محمد

أستاذ التعليم العالي  
جامعة ابن خلدون، تيارت

## الإحصاء الرياضي والنماذج الإحصائية سلسلة دروس مع تمارين مختارة

موجه إلى الشعب التالية:

- ◀ شعب علوم الطبيعة والحياة، علوم الأرض والكون
- ◀ شعب العلوم الإنسانية (العلوم التجارية وعلوم التسيير،  
علم الاجتماع، علم النفس...)

إعادة الطبعة الأولى



ديوان المصطبوعات الجامعية

© ديوان المطبوعات الجامعية: 2015-02  
رقم النشر: 4.01.4918  
رقم ر.د.م.ك (ISBN): 978.9961.0.1101.0  
رقم الإيداع القانوني: 2007- 4156

# إهداء

إلى روح والدي الحاج المختار رحمه الله .



## مقدمة الطبعة الثانية

إن المكتبة العربية تفتقر كثيرا إلى الكتب العلمية في مختلف التخصصات التطبيقية والتكنولوجية، كما أن الجامعة الجزائرية تعتمد كليا على المراجع الأجنبية في هذا النوع من العلوم، الأمر الذي خلق صعوبات لدى طلبة الجيل الحالي الذي أمضى جزءا غير يسير من مشواره الدراسية يدرس باللغة العربية. ومن منطلق أن الإحصاء أصبح يلعب دورا أساسيا في كل العلوم المختلفة، كعلوم التسيير والإجتماع وعلم الأحياء والطب وغيرها...، فقد جاء هذا الكتاب كأداة مساعدة للجميع، بأسلوب سهل ومختصر، يعتمد على أكثر عدد من الأمثلة والتمارين، وبعيدا عن العمليات الرياضية ذات البرهنة المعقدة.

إن الهدف من هذا المرجع هو تقديم الأسس الأولية الضرورية التي تفيد كل فرد حسب تخصصه، متبعا كل المقررات التي تدرس حاليا في الجامعة الجزائرية، في التخصصات السابقة الذكر، على أن الرموز المستعملة فيه هي رموز لاتينية، الهدف منها هو توحيد المصطلحات العالمية، تسهيلا للطلاب لمواصلة البحث مع لغات أخرى.

كنت هو الحال في الطبعة الأولى، فقد شمل هذا العمل ستة فصول، منقحة مع إضافة بعض التعاريف لبعض القوانين، كما أن كل فصل قد تدعم بمجموعة من الأمثلة و التمارين المحلولة، مع إضافة مجموعة أخرى من المسائل والتمارين الغير المحلولة في نهاية كل فصل، حتى نتيح الفرصة للطلاب لإختبار قدراته الإستيعابية.

جاء الفصل الأول كمقدمة لنظرية الاحتمال، به قسم خاص بالمتغيرات العشوائية والتوزيعات الاحتمالية؛ وتناول الفصل الثاني، الثالث والرابع إختبار الفرضيات، نظرية العينات ونظرية التقدير على التوالي، كمقدمة لتحليل البيانات الإحصائية، وهي تفيد بشكل أكبر، الذين يعتمدون على التجارب

العلمية، وتحليل نتائجها؛ أما الفصل الخامس فقد تناول نظرية الارتباط، التي تعد ذات أهمية قصوى في حد ذاتها، وهي بذلك تكون مقدمة لجميع الذين يبحثون كيفية بناء نماذج رياضية، ذات متغير واحد (الارتباط البسيط)، أو ذات عدة متغيرات (الارتباط متعدد)؛ وكانت خاتمة الكتاب بفصل سادس خاص بالسلاسل الزمنية وتحليلها، نظرا لكون الكثير من البيانات الإحصائية تعطى بدلالة الزمن، خاصة تلك التي تدرس بعض الظواهر الاجتماعية والمؤشرات الاقتصادية.

نأمل أن يكون هذا العمل البسيط، ثمرة طيبة بين أيدي طلبتنا الكرام، وأن يحقق الفائدة المرجوة منه، وأن يكون فاتحة لأعمال أخرى وصدقة جارية للذي سهر من أجل تعليمي، والذي الحاج المختار (رحمة الله عليه) وللشهداء الأبرار الذين سقطوا من أجل أن نحيا نحن في هذا الوطن الغالي أحرارا شرفاء.

الأستاذ: م. معتوق



# الفصل الأول

## نظرية الاحتمالات

### مقدمة:

نظرية الاحتمالات هي قسم من الرياضيات، يدرس الاتجاه العام للظواهر العرضية (العشوائية). تلعب هذه النظرية دورا مهما في مختلف العلوم، خاصة في العلوم الاقتصادية وعلوم الطبيعة والحياة، حيث يتم من خلالها دراسة مختلف الظواهر العرضية التي تتكرر والتي تستوجب وضع قوانين يمكن الحكم على صحته أو خطئه.

### 1.1- مبادئ الحساب الاحتمالي:

#### 1.1.1- مفاهيم أساسية:

- التجربة: تعد من أهم المفاهيم في نظرية الاحتمالات، فهي تقوم على أساس التأكد من تحقق وصول بعض الظروف المشتركة لتجربة ما (ظروف من صنع الإنسان أو ظروف وليدة الصدفة). مثل رمي زهرة نرد، دراسة كمية التساقط في منطقة ما...

وتنقسم التجربة إلى تجربة نظامية، حيث أنه يمكن توقع نتائجها سلف على أساس قوانين علمية معروفة، إنطلاقا من جملة من الشروط. والتجربة العشوائية، وهي التجربة التي يمكن تكرارها، على أن نتائجها غير محددة سلفا، فهي تعتمد على الصدفة رغم إنطلاقها من نفس الشروط.

- فراغ العينة: وهي مجموعة النتائج الممكنة الكلية لتجربة ما، ويمكن أن يكون غير منته (عدد غير محدود من الإمكانيات)، ومنته (عدد غير محدود من الإمكانيات). يرمز لفراغ العينة بالرمز  $\Omega$ .

- الحدث و أنواعه: إن النتيجة أو النتائج المحددة من النتائج الممكنة لتجربة ما تشكل حدثا. وينقسم الحدث إلى: حدث بسيط، غير قابل للتجزئة، فهو مجموعة جزئية من فراغ العينة؛ حدث مركب، أمكن تفكيكه إلى حوادث بسيطة؛ حدث أكيد وهو حدث مؤكد يحوي جميع الحوادث البسيطة؛ حدث مستحيل، وهو حدث غير قابل للتحقق، أي أنه مجموعة خالية من المجموعة الكلية التي تمثل فراغ العينة؛ حدث متمم أو معاكس وهو حدث  $\bar{A}$  مرتبط بالمجموعة  $\Omega$ ، يتكون من مجموعة الإمكانيات الغير محققة لـ  $A$ ، حيث:

$$\bar{A} = \Omega - A \quad (1.1)$$

حوادث متنافية و هي حوادث لا يمكن وقوعها في أن واحد ووقوع أحدها يمنع وقوع الآخر؛ حوادث غير متنافية، وهي عكس المتنافية، فوقوع  $A$  لا يمنع وقوع الحدث  $B$ ؛ حوادث مستقلة و هي حوادث لا يؤثر وقوع أحدهما على الآخر؛ حوادث غير مستقلة أي مرتبطة وهي حوادث شرطية، فوقوع الحدث  $A$  يشترط وقوع الحدث  $B$ .

### 2.1.1- تعريف الإحتمال:

إذا كان  $m$  هو عدد الحالات الملائمة للحدث  $A$  ؛ وكان  $n$  هو عدد حالاته الممكنة، فإن إحتمال وقوع الحدث  $A$  يكون وفق العلاقة التالية:

$$P(A) = \frac{m}{n} \quad \text{حيث} \quad 1 \leq m \leq n \quad (1.2)$$

➤ مثال 1:

ترمى قطعة نرد مرة واحدة، نجد أن فراغ العينة (الحالات الممكنة) لهذه التجربة هو:

$$\Omega = \{1, 2, 3, 4, 5, 6\} \quad \text{أي أن } n = 6$$



ولحساب احتمال الحصول على عدد زوجي، فإن الحالات الملائمة هي:

$$P = \{2, 4, 6\} \quad \text{أي أن } m = 3$$

$$P(A) = \frac{m}{n} = \frac{3}{6} = 0.5 \quad \text{إذن:}$$

➤ مثال 2:

يحتوي كيس على 10 كرات، 7 منها حمراء اللون. أحسب احتمال الحصول على 4 كرات حمراء إذا سحبنا من الكيس 6 كرات.

إن احتمال الحصول على 4 كرات حمراء من الكيس بعد سحب 6 كرات هو:

$m$ : عدد الحالات الملائمة وهو سحب 6 من 10 أي:

$$C_{10}^6 = \frac{10!}{(10-6)! \times 6!} = 210$$

$n$ : عدد الحالات الملائمة وهو سحب 6 من 10 أي:  $C_7^4 \times C_3^2 = 105$

إذن:

$$P(A) = \frac{m}{n} = \frac{105}{210} = 0.5$$

❖ صيغة أخرى لتعريف الإحتمال:

إن احتمال تحقق الحدث  $A$  المرتبط بتجربة عشوائية ما هو نهاية التكرار النسبي للحدث  $A$  وذلك عندما تتكرر هذه التجربة في عدد كبير من المرات، أي أن:

$$p(A) = \lim_{n \rightarrow \infty} \frac{m}{n} \quad (1.3)$$

$$n \rightarrow \infty$$

### 3.1.1- خواص الإحتمال:

إذا كان  $A$  حدث ما، فإن:

$$0 \leq P(A) \leq 1$$

• إذا كان  $\bar{A}$  حدث متمم لـ  $A$ ، فإن:

$$p(A) + p(\bar{A}) = p(A \cup \bar{A}) = p(\Omega) = 1$$

$$p(\bar{A}) = 1 - p(A)$$

ومنه:

### 4.1.1- الخواص الأساسية في نظرية الإحتمال:

(أ) - قاعدة الجمع للأحداث المتنافية:

إذا كان  $A$  و  $B$  حدثين متنافيين، فإن تحقق مجموعهما هو مجموع احتمالهـما، أي أن:

$$P(A \cup B) = P(A) + P(B) \quad (1.4)$$

وبصفة عامة إذا كان لدينا مجموعة من الحوادث  $A_1, A_2, \dots, A_n$

فإن:

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

أي أن:

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) \quad (1.5)$$

➤ مثال:

ترمى زهرة نرد مرة واحدة. ما هو احتمال الحصول على الرقم 1 أو الرقم 2.

إن احتمال الحصول على الرقم 1 هو:  $P(A) = \frac{m}{n} = \frac{1}{6}$

وإحتمال الحصول على الرقم 2 هو:  $P(B) = \frac{m}{n} = \frac{1}{6}$

وعليه هو احتمال الحصول على الرقم 1 أو الرقم 2 هو:

$$p(A \cup B) = P(A) + P(B) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

(ب) - قاعدة الجمع للأحداث الغير المتنافية:

إذا كان  $A$  و  $B$  حدثين غير متنافيين، فإن تحقق مجموعتهما هو مجموع احتماليهما ولكن نستثني احتمال وقوعهما معا في آن واحد، أي أن:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (1.6)$$

ويمكن تعميم هذه القاعدة على أكثر من حدثين، فلو كان لدينا ثلاثة حوادث  $A$  و  $B$  و  $C$  فإن:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

ويمكن البرهنة على ذلك بسهولة، وذلك بفرض أن  $B \cup C$  هي مجموعة واحدة ولتكن  $D$  ومن ثم تطبيق القاعدة (1.6)، مع ملاحظة أن عملية  $U$  هي توزيعية بالنسبة لـ  $\cap$ .

وبصفة عامة إذا كان لدينا أكثر من ثلاثة حوادث فإن صيغة الجمع للأحداث الغير المتنافية تأخذ الشكل التالي:

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum p(A_j \cap A_k) + \sum P(A_i \cap A_j \cap A_k) - \dots \dots \dots (1.7)$$
$$\dots \dots \dots + (-1)^{n-1} P(A_1 \cap A_2 \dots \cap A_k)$$

(ج) - قاعدة الضرب للأحداث المستقلة:

إن احتمال وقوع الحدثين  $A$  و  $B$  في آن واحد، مع العلم أنهما مستقلين، هو حاصل ضرب احتمال وقوع كل منهما، أي أن:

$$P(A \cap B) = P(A).P(B) \quad (1.8)$$

وبصفة عامة، إذا كان لدينا أكثر من حدثين مستقلين مثني مثني، فإن القاعدة (1.7) تصبح على الشكل التالي:

$$P\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P(A_i) \quad (1.9)$$

➤ **تمارين:**

يحتوي صندوق على 10 مصابيح كهربائية، 7 منها صالحة و 3 غير صالحة. سحبنا عشوائيا من الصندوق مصباحين (سحب مع الإرجاع للمصباح الأول). أحسب ما يلي:

- احتمال المصباحين صالحين،
- احتمال المصباحين غير صالحين،
- احتمال الأول صالح و الثاني غير صالح.

➤ **الحل:**

(أ) - احتمال المصباحين صالحين، يعني المصباح الأول صالح A والثاني صالح B:

$$P(A \cap B) = P(A) \times P(B) = \frac{7}{10} \times \frac{7}{10} = \frac{49}{100} = 0.49$$

$$\text{لأن: } P(A) = \frac{7}{10} \text{ و } P(B) = \frac{7}{10}$$

(ب) - احتمال المصباحين غير صالحين، يعني الأول غير صالح C و الثاني غير صالح D:

$$P(C \cap D) = P(C) \times P(D) = \frac{3}{10} \times \frac{3}{10} = \frac{9}{100} = 0.09$$

(ج) - احتمال الأول صالح و الثاني غير صالح، يعني الأول صالح A و الثاني غير صالح C:

$$P(A \cap C) = P(A) \times P(C) = \frac{7}{10} \times \frac{3}{10} = \frac{21}{100} = 0.21$$

(د) - قاعدة الضرب للأحداث المرتبطة (الإحتمال الشرطي):

إن تحقق الحدث  $A$  يكون مرتبط بتحقق الحدث  $B$  بصورة مسبقة، أي أن تحقق  $A$  يشترط تحقق  $B$ . نسمى في هذه الحالة حساب وقوع  $A$  بشرط  $B$  بالإحتمال الشرطي و نكتب:

$$P(B) \neq 0 \text{ بحيث } P(A/B) = \frac{P(A \cap B)}{P(B)} \quad (1.10)$$

$$P(A) \neq 0 \text{ بحيث } P(B/A) = \frac{P(A \cap B)}{P(A)} \quad (1.11) \quad \text{أو}$$

➤ تمرين:

نسبة الرسوب، حسب محاضر الامتحانات النهائية لمعهد ما، هو 15% في مقياس الإحصاء؛ وفي مقياس الرياضيات هو 25% ؛ وفي المقياسين معا هو 10%. أخذنا طالبا عشوائيا من المحاضر. إذا كان هذا الطالب راسبا في الرياضيات ما هو إحتمال أن يرسب في الإحصاء؟. إذا كان هذا الطالب راسبا في الإحصاء ما هو إحتمال أن يرسب في الرياضيات؟. ماهو إحتمال أن يرسب في الرياضيات أو في الإحصاء؟.

➤ الحل:

إحتمال أن يرسب في الرياضيات هو  $P(A) = 0.15$

إحتمال أن يرسب في الإحصاء هو  $P(B) = 0.25$

إحتمال أن يرسب في الرياضيات و الإحصاء معا هو

$$P(A \cap B) = 0.10$$

• إحتمال أن يرسب في الإحصاء مع العلم أنه (أي بشرط) راسبا في الرياضيات هو:

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{0.10}{0.15}$$

- احتمال أن يرسب في الرياضيات مع العلم أنه راسبا في الإحصاء هو:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{0.10}{0.25}$$

- احتمال أن يرسب في الرياضيات أو في الإحصاء هو:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.15 + 0.25 - 0.10 = 0.30$$

### هـ) - نظرية بايز:

تظهر نظرية بايز كيفية حساب الاحتمالات الشرطية لحوادث متنافية تشكل مجموعة كلية و مرافقة لحدوث الحدث  $A$ .

لنفترض أنه لدينا مجموعة من الحوادث المتنافية  $B_1, B_2, \dots, B_n$  مشكلة مجموعات جزئية من المجموعة الكلية  $\Omega$ ؛ و لنفترض الحدث  $A$  و الذي يتحقق بشرط تحقق أحد الحوادث السابقة  $B_1, B_2, \dots, B_n$ .

نريد حساب الاحتمالات الشرطية  $P(B_i/A)$  وذلك بمعلومية  $P(B_i)$ ، أي بشرط تحقق  $P(B_i)$ ، فيكون لدينا:  
من جهة ثانية لدينا:

$$P(B_i/A) = \frac{P(A \cap B_i)}{P(A)} \Rightarrow P(A \cap B_i) = P(B_i/A) \times P(A) \dots\dots 2$$

وحيث أن الطرفين 1 و 2 متساويان، يكون لدينا:

$$P(A \cap B_i) = P(A) \times P(B_i/A) = P(B_i) \times P(A/B_i)$$

بالقسمة على  $P(A)$  حيث  $P(A) \neq 0$  نحصل على:

$$P(B_i/A) = \frac{P(B_i) \times P(A/B_i)}{P(A)} \quad (*)$$



الحدث A لا يتحقق إلا بتحقق أحد هذه الحوادث، فإن:

$$P(A) = \sum P(B_i) \times P(A/B_i)$$

وبالتعويض في المعادلة (\*) نجد:

$$P(B_i/A) = \frac{P(B_i) \times P(A/B_i)}{\sum P(B_i) \times P(A/B_i)} \quad (1.12)$$

وهي العلاقة النهائية لنظرية بايز.

➤ **تمرين:**

تملك مؤسسة إنتاجية ثلاث آلات للإنتاج، بنسب إنتاج هي على التوالي: 60%، 30%، 10% من إجمالي إنتاج المؤسسة. فإذا كانت نسبة الإنتاج الصالح لهذه الآلات هي على التوالي: 98%، 97%، 96%. سحبنا بصورة عشوائية وحدة من المؤسسة، ووجد بأنها فاسدة، فما هو احتمال أن تكون هذه الوحدة من إنتاج الآلة الثالثة.

➤ **الحل:**

نسمي الآلات الثلاث على التوالي A، B و C وهي ثلاث مجموعات جزئية متنافية، إتحادهما يشكل المجموعة الكلية (المؤسسة) و نرسم للوحدة الفاسدة بالرمز X والتي تتحقق بتحقق إحدى المجموعات السابقة، و المطلوب هو إيجاد:

احتمال أن تكون هذه الوحدة الفاسدة من إنتاج الآلة C، و بتطبيق نظرية بايز نجد:

$$P(C/X) = \frac{P(C) \times P(X/C)}{P(A) \times P(X/A) + P(B) \times P(X/B) + P(C) \times P(X/C)}$$

بالتطبيق العددي نجد:

$$P(C/X) = \frac{0.10 \times 0.04}{(0.6 \times 0.02) + (0.30 \times 0.03) + (0.10 \times 0.04)} = 0.16$$

## 2.1- المتغيرات العشوائية و التوزيعات الإحتمالية:

### 1.2.1- المتغيرات العشوائية:

➤ مثال:

ترمى زهرة نرد مرتين على التوالي. ليكن  $X$  المتغير العشوائي الذي يمثل مجموع العددين الظاهرين، أي:

$X(a,b) = a+b$ . المطلوب شكل التوزيع الإحتمالي المناسب، ثم مثله بيانيا.

➤ الحل:

لنشكل فضاء العينة، أي كل الحالات الكلية الممكنة الناتجة عن هذه التجربة الإحتمالية:

$\Omega = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}$ .

إن قيم  $X$  الموافقة للقانون  $X(a,b) = a + b$  هي كما يلي:

$$X = \{ 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 \}$$

إن إحتمال أن يكون:

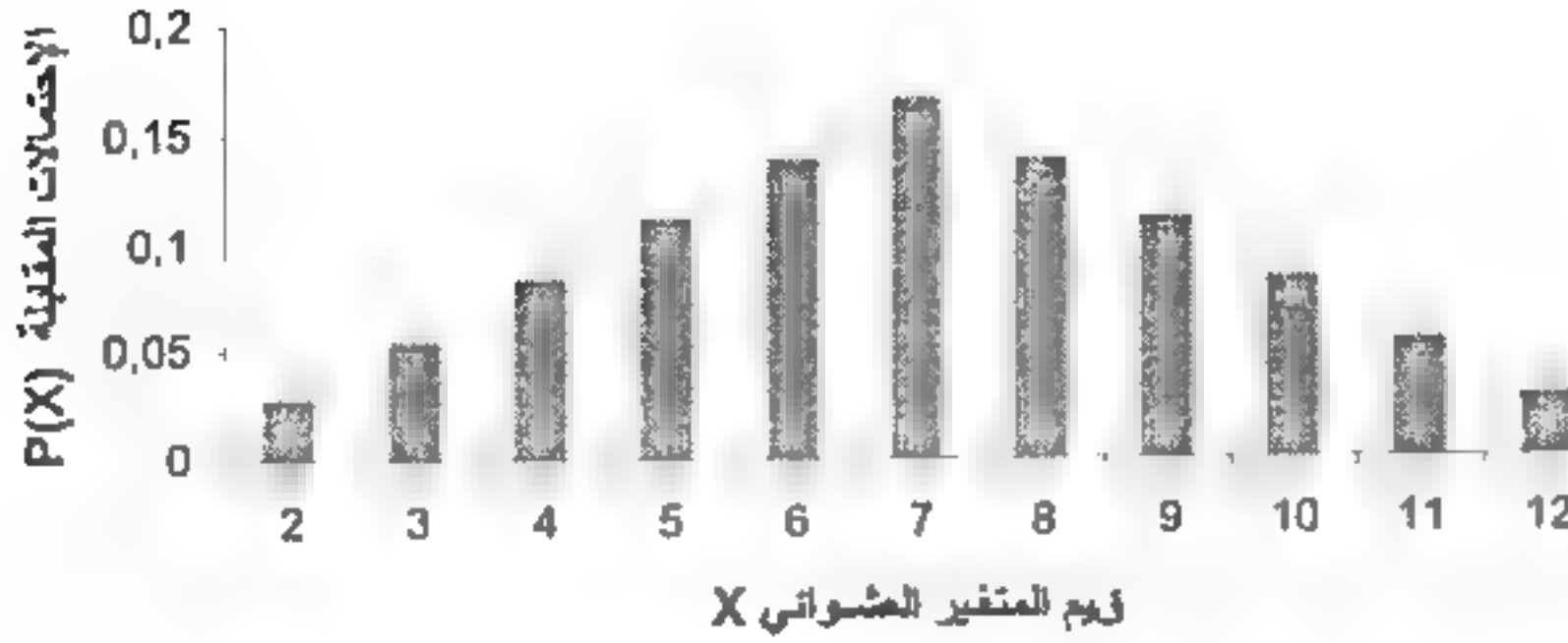
$$P(x=2) = \frac{1}{36}$$

لأنه توجد حالة واحدة ممكنة يكون فيها مجموع العددين الظاهرين يساوي 2 وهي (1,1) من بين 36 حالة كلية. وبالمثل لباقي قيم  $X$ ؛ فمثلا  $P(x=3) = \frac{2}{36}$  التي تقابل الحالتين (1,2) و (2,1)، وهكذا.....

قيم $X$	2	3	4	5	6	7	8	9	10	11	12
$P(X)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

ويمكن بسهولة تمثيل هذا التوزيع بيانياً، فنحصل على مايلي:

التمثل البياني بواسطة مدرج للتوزيع الإحتمالي



نلاحظ أن هذا التوزيع متماثل حول نقطة هي 7.

من هذا المثال البسيط يمكن أن نصل إلى تعريف المتغير العشوائي. إن المتغير العشوائي  $X$  هو مجموعة قيم لنتائج تجربة إحصائية ما، مقترن باحتمالات مقابلة، فلكل قيمة  $X_i$  احتمال مقابل  $P(X_i)$ ، وهو ما يشكل توزيعاً احتمالياً.

#### 1.1.2.1- أنواع المتغيرات العشوائية:

و يمكن أن نميز: المتغير العشوائي المنفصل والمتغير العشوائي المتصل.

##### (أ) - المتغير العشوائي المنفصل:

وهو المتغير العشوائي الذي يمكنه أخذ عدداً منتهياً من القيم الصحيحة ضمن مجال تغيره، و يدعى في هذه الحالة بالمتغير العشوائي المنفصل ضمن فراغ عينة منته، مثل رمي حجر نرد مرة واحدة. أما المتغير العشوائي المنفصل ضمن فراغ عينة غير منته ولكنه قابل للعد، فهو العشوائي الذي يمكنه أخذ عدداً غير منتهياً من القيم الصحيحة ضمن مجال تغيره ألا أن يكون هذا المجال قابلاً للعد. مثال على ذلك عمليات السحب بالإعادة، حيث أن عدد مفردات السحب يبقى دائماً ثابتاً، وأن عملية سحب ما هي إلا متغير عشوائي مستقل عن العمليات الأخرى.

لنفرض أن  $X$  هو متغير عشوائي منفصل، و أن  $\Omega$  هو فراغ عينته، فيكون لدينا:

$$\Omega = \{X_1, X_2, \dots, X_n\}$$

وأن الإحتمالات المقابلة لقيم  $X_i$  هي  $P(X=X_i)$ ، فإن التوزيع الاحتمالي يكون كما يلي:

قيم $X$	$X_1$	$X_2$	.....	$X_n$
$P(X)$	$P(X_1)$	$P(X_1)$	.....	$P(X_n)$

إن هذا التوزيع يجب أن يحقق شرطين أساسين:

$$\begin{cases} P(X_i) \geq 0 \\ \sum_{i=1}^n P(X_i) = 1 \end{cases} \quad (1.13)$$

➤ مثال:

ترمى قطعة نقود مرتين على التوالي. نعرف المتغير العشوائي بأنه عدد مرات الحصول على الوجه  $F$ . و المطلوب شكل التوزيع الاحتمالي المناسب لهذه التربة الاحتمالية مع الرسم البياني.

➤ الحل:

نرمز بـ  $F$  للوجه الأول لقطعة النقود؛ وبالرمز  $P$  للوجه الآخر، فيكون فضاء العينة كما يلي:

$$\Omega = \{(F,F), (F,P), (P,F), (P,P)\}.$$

إن قيم  $X$  المقابلة هي كما يلي:

$X = 0$  عدم الحصول على الوجه  $F$ ، أي الحصول على الوجه  $(P,P)$

$X = 1$  الحصول على الوجه  $F$  مرة واحدة، أي  $(F,P), (P,F)$

$X = 2$  الحصول على الوجه  $F$  مرتين، أي الحصول على الوجه  $(F,F)$   
أي أن:  $X = \{ 0, 1, 2, \}$

أما الإحتمالات المقابلة فهي:

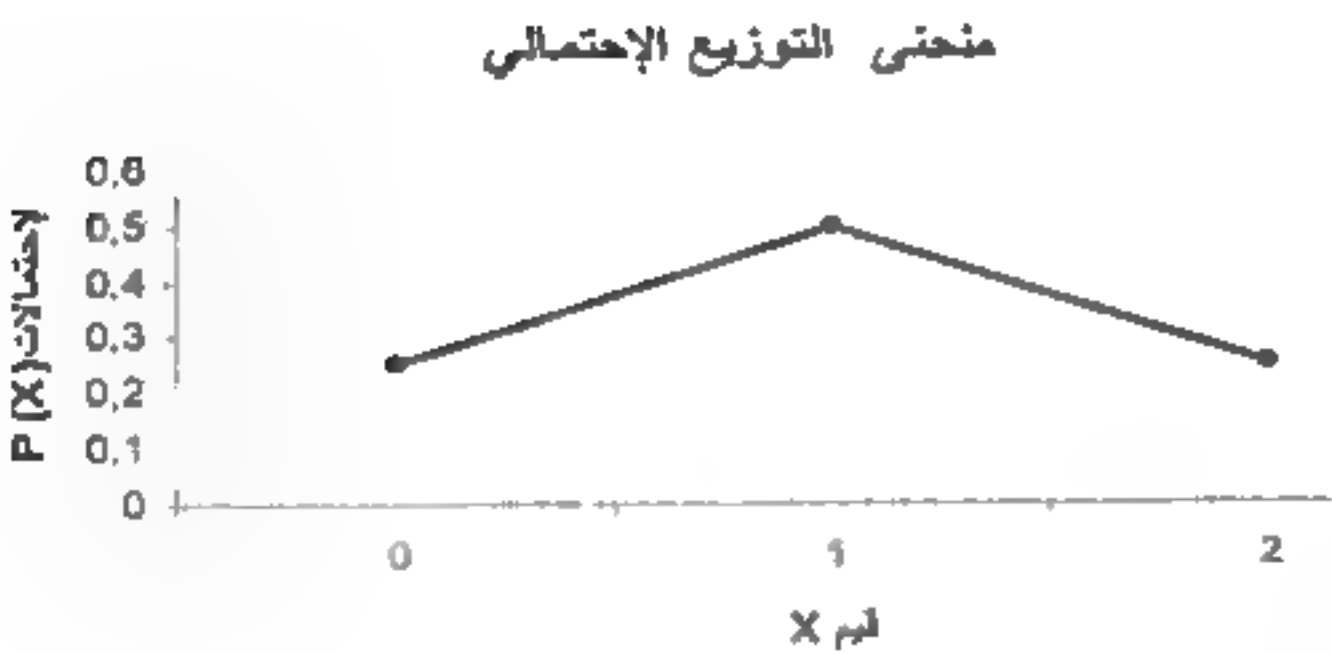
$P(x=2)=\frac{1}{4}$  و  $P(x=1)=\frac{2}{4}=\frac{1}{2}$  و  $P(x=0)=\frac{1}{4}$   
نلخص هذه النتائج في الجدول التالي:

المجموع	2	1	0	قيم $X$
1	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	$P(X)$

نلاحظ أن:

$$\begin{cases} P(X_i) \geq 0 \\ \sum_{i=1}^n P(X_i) = 1 \end{cases}$$

أما التمثيل البياني لهذا التوزيع فهو كما يلي:



(ب) المتغير العشوائي المتصل:

كون  $X$  متغيرا عشوائيا متصلا، إذا كان بإمكان هذا المتغير أن يأخذ أية قيمة في مجال تغير  $X$ ؛ إن إحتمال أن يأخذ  $X$  قيمة واحدة صحيحة يكون

معدوما في هذه الحالة (عكس المتغير العشوائي المنفصل). فمثلا أن طول شخص ما هو متغير عشوائي متصل، ذلك أن هذا الشخص يمكن أن يكون طوله 1.80، وبصورة أدق يمكن أن يكون 1.801 ويمكن أن يكون 1.8015 إذا إستعملنا وسائل قياس أدق وهكذا....، فاحتمال أن يكون طول هذا الشخص مساويا بالضبط 1.80 هو احتمال معدوم.

إن قانون المتغير العشوائي المتصل يأخذ شكل الدالة المستمرة في مجال تعريفها، فتكون على الشكل التالي:

$$f(x) = \begin{cases} f(x) & a \leq x \leq b \\ 0 & \text{دون ذلك} \end{cases}$$

وتسمى الدالة  $f(x)$  بدالة الكثافة الإحتمالية، عندما تعبر عن احتمال قيمة معينة كما يلي  $P(X = x)$  ونكتب أيضا:  $f(x)$  وتسمى الدالة  $f(x)$  دالة الكثافة الاحتمالية. ولكي يمكن اعتبار دالة ما، أيا كانت، دالة كثافة احتمالية يجب أن يتحقق شرطان اثنان:

- 1)  $f(x) \geq 0$
- 2)  $\sum_x f(x) = 1$

وذلك في حالة المتغيرات العشوائية المتقطعة.

أما في حالة المتغيرات العشوائية المستمرة فهي مجموعة القيم اللامتناهية، في مجال محدود، التي يمكن أن يأخذها المتغير  $X$  والاحتمالات الملحقه بها. ولكي تكون  $f(x)$  دالة كثافة إحتمالية، يجب تحقق الشرطين:

$$\begin{cases} f(x) \geq 0 \\ \int_a^b f(x)dx = 1 \end{cases} \quad (1.14)$$

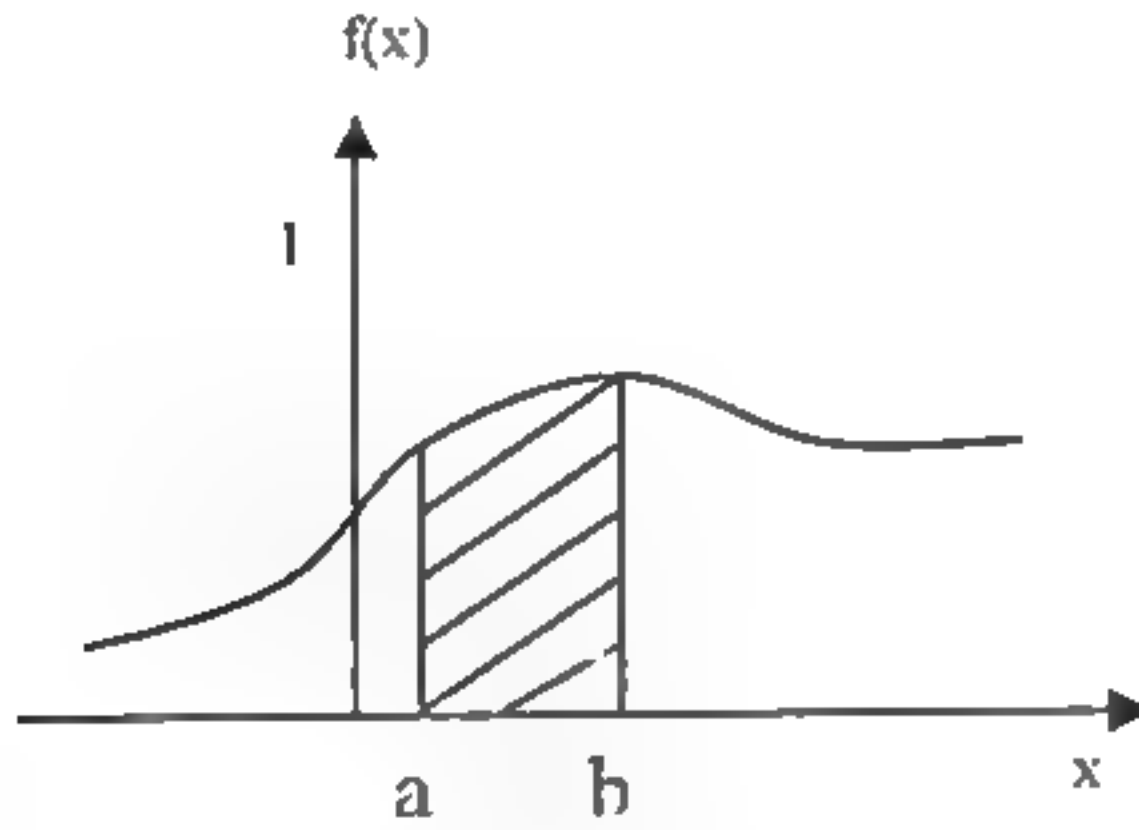


وهي ممثلة بمنحنى متصل.

لاحظ أنه بما أن  $X$  تأخذ عددا لا متناهيا من القيم داخل أي مجال، فإن احتمال قيمة بعينها يكون ذلك بحساب المساحة تحت منحنى  $f(x)$  بين حدود المجال.

$$P(a < x < b) = \int_a^b f(x)dx \quad (1.15)$$

لاحظ أن إشارة التكامل هنا تقابل إشارة المجموع في حالة المتغير العشوائي المتقطع.



التمثيل البياني لدالة الكثافة للمتغيرة العشوائية المستمرة

➤ مثال لحالة المتغيرات العشوائية المتقطعة

نأخذ دالة الكثافة لـ  $X$  نتيجة لإلقاء حجر نرد:

$$f(1) = f(2) = f(3) = \dots f(6) = 1/6 \geq 0, \quad \text{الشرط الأول محقق}$$

والشرط الثاني أيضا لأن:

$$\sum f(x) = 1/6 + 1/6 + \dots + 1/6 = 6(1/6) = 1$$

➤ مثال لحالة المتغيرات العشوائية المستمرة

ليكن  $X$  متغيرا عشوائيا متصلا معرفا بالدالة الإحتمالية التالية:

$$f(x) = \begin{cases} KX & 1 \leq x \leq 3 \\ 0 & \text{دون ذلك} \end{cases}$$

- أوجد قيمة  $K$  حتى تكون  $f(x)$  دالة كثافة إحتمالية،
- أوجد  $P(1 \leq x \leq 2)$ ،
- مثل بيانيا هذه الدالة.

➤ الحل:

حتى تكون  $f(x)$  دالة كثافة إحتمالية يجب أن يكون:

$$\int_1^3 f(x)dx = 1 \Rightarrow \int_1^3 kx dx = 1 \Rightarrow k \int_1^3 x dx = 1$$

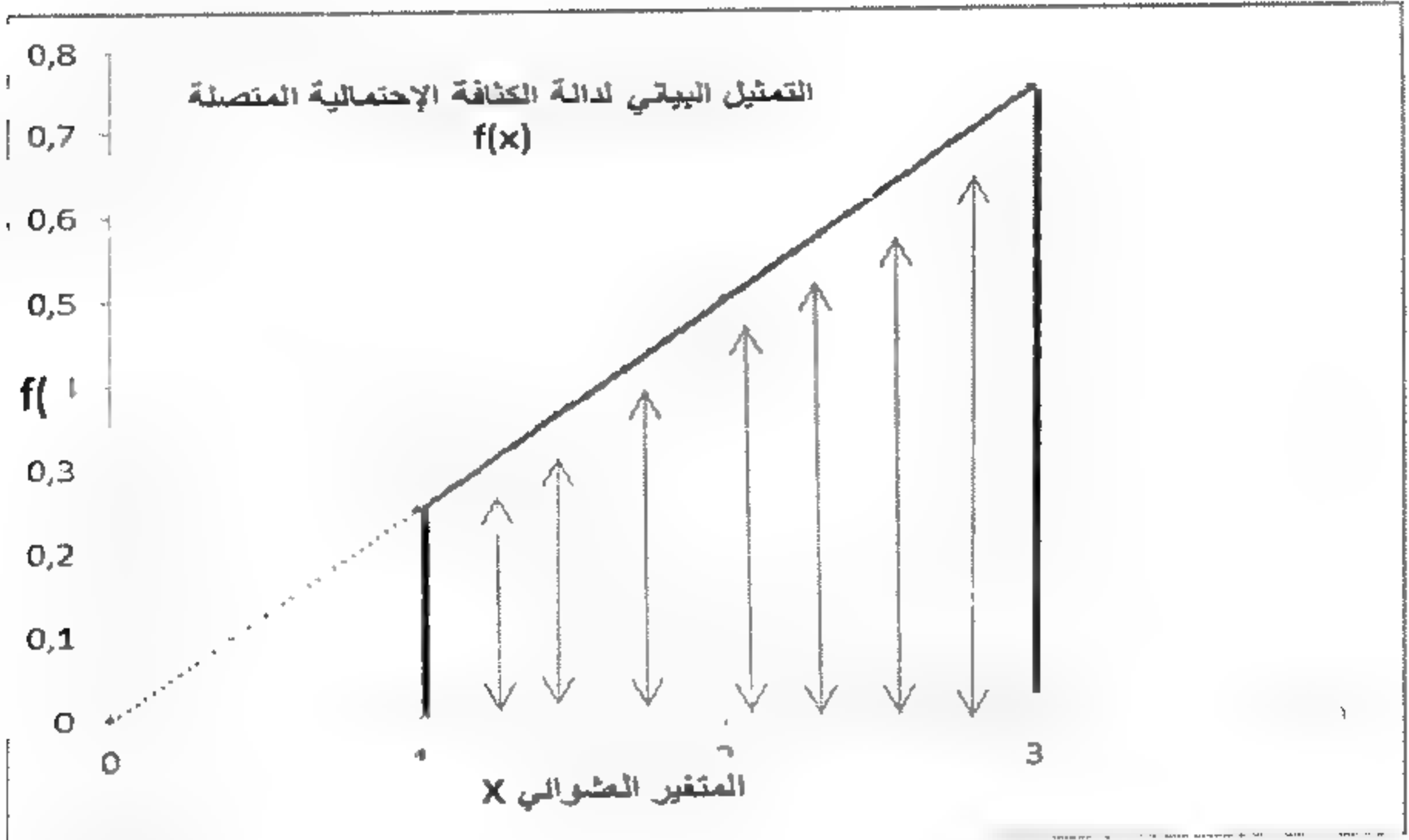
$$k \left[ \frac{x^2}{2} \right]_1^3 = k \left[ \frac{9}{2} - \frac{1}{2} \right] = 1 \Leftrightarrow k = \frac{1}{4}$$

وتكون الدالة على النحو التالي:

$$f(x) = \begin{cases} 1/4X & 1 \leq x \leq 3 \\ 0 & \text{دون ذلك} \end{cases}$$

ويمكن بسهولة إيجاد:

$$P(1 \leq x \leq 2) = \int_1^2 \frac{1}{4}x dx = \frac{1}{4} \left[ \frac{x^2}{2} \right]_1^2 = \frac{1}{4} \left[ \frac{4}{2} - \frac{1}{2} \right] = \frac{3}{8} = 0.375$$



ملاحظة هامة:

يمكن إيجاد  $P(1 \leq X \leq 2)$  هندسيا وذلك بحساب مساحة شبه المنحرف  
وفق العلاقة:  $\frac{Ax+B}{2} \times h$

$$P(1 \leq x \leq 2) = \frac{0.5 + 0.25}{2} \times 1 = 0.375 \quad \text{أي أن:}$$

● دالة التوزيع  $F(x)$  للمتغيرة العشوائية المتقطعة

تعرف دالة التوزيع - وتسمى أيضا "الدالة التجميعية" - كما يلي:

$$F(x) = P(X \leq x)$$

ويمكن استنتاج دالة التوزيع من دالة الكثافة الاحتمالية  $f(x)$  كما يلي:

$$F(x) = P(X \leq x) = \sum_{u \leq x} f(u)$$

إذا كانت  $X$  تأخذ عددا منتهيا من القيم فإن  $F(x)$  يمكن تعريفها كما يلي:

$$F(x) = \begin{cases} 0 & , & -\infty < x < x_1 \\ f(x_1) & , & x_1 \leq x < x_2 \\ f(x_1) + f(x_2) & , & x_2 \leq x < x_3 \\ \dots & \\ f(x_1) + f(x_2) + \dots + f(x_n), & x_n \leq x < +\infty \end{cases}$$

➤ تمرين

في تجربة رمي قطعة نقدية مرتين نعتبر المتغير العشوائي  $X$  التي تمثل عدد مرات الحصول على كتابة. إن القيم الممكنة لـ  $X$  هي 0، 1، 2. في نفس التجربة نغرق المتغير العشوائي  $Y$  الذي يمثل عدد مرات الحصول على صورة، فتكون القيم الممكنة 0، 1، 2، ونعرف المتغير  $Z = X - Y$  بحث  $Z$  فتكون القيم الممكنة له هي 0، 2، -2. الاحتمالات الملحقة بقيمها يمكن حسابها كما يلي:

$$P(Z = 0) = P(X - Y = 0) = P(X = 0 \text{ et } Y = 0 \text{ ou } X = 1 \text{ et } Y = 1 \text{ ou } X = 2 \text{ et } Y = 2)$$

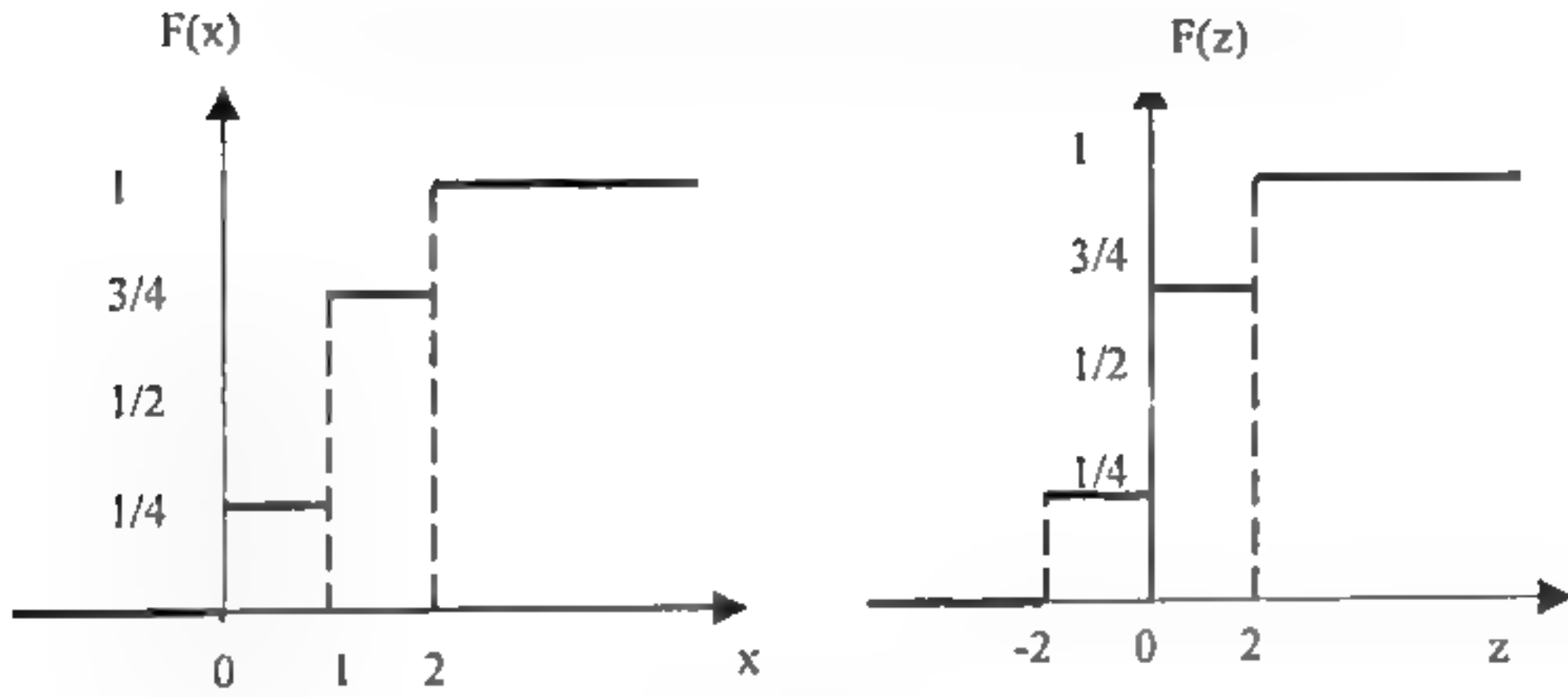
$$\Rightarrow P(Z = 0) = 0 + P(X = 1 \text{ et } Y = 1) + 0 = 2 * 0.5^2 = 0.5$$

ويمكن إيجاد قيم  $F(x)$  و  $F(z)$  للمتغيرين  $X$  و  $Z$  كما في الجدولين:

$X$	0	1	2
$f(x)$	1/4	1/2	1/4
$F(x)=P(X \leq x)$	1/4	3/4	1

$Z$	-2	0	2
$f(x)$	1/4	1/2	1/4
$F(x)=P(X \leq x)$	1/4	3/4	1

## التمثيل البياني لدالة التوزيع للمتغيرة العشوائية



● دالة التوزيع  $F(x)$  للمتغير العشوائي المستمر

تعرف دالة التوزيع للمتغير المستمر كما يلي:

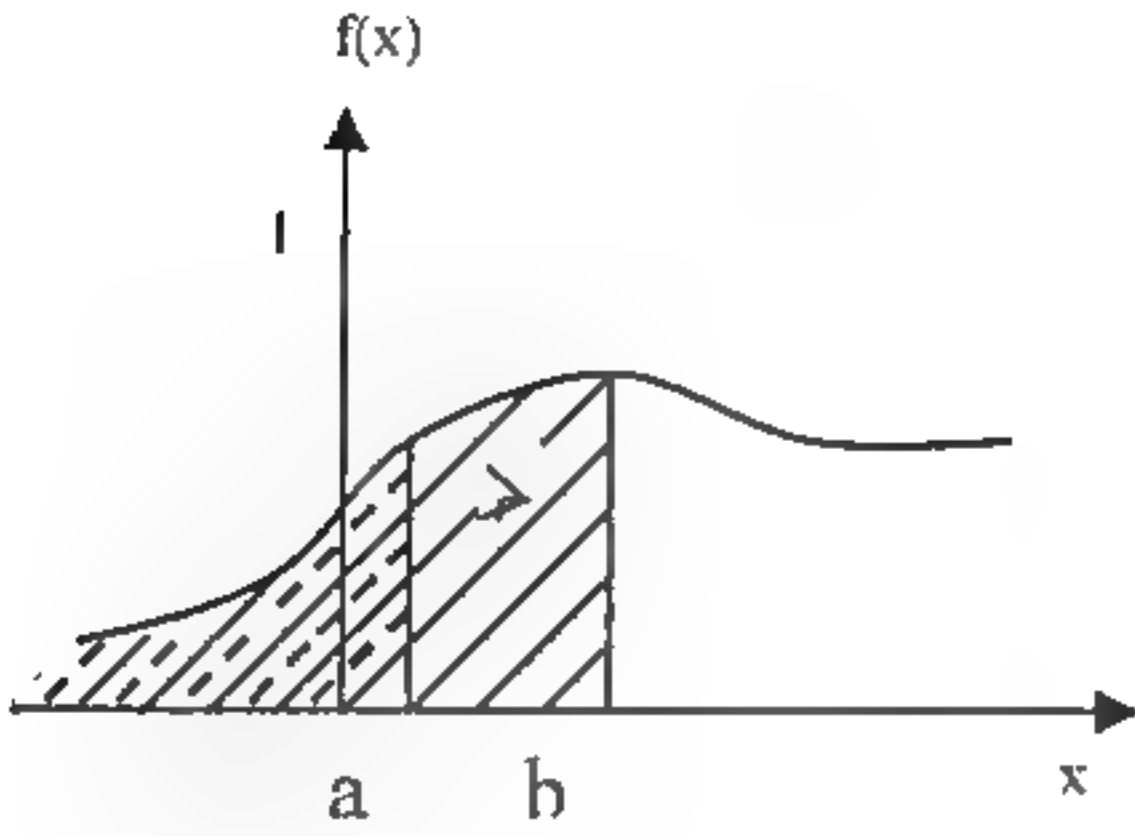
تسمى  $F(X)$  دالة التوزيع للمتغير المستمر  $X$  إذا كان:

(1.16)

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u) du$$

ولحساب احتمال جال من الجهة اليسرى، فإننا نقوم بالتعويض في دالة التوزيع بدلا من حساب التكامل وفق القاعدة التالية: بفرض  $a, b$  نقطتان من مجال تعريف  $X$ ، بحيث  $b > a$ . لحساب احتمال أن تكون  $X$  تنتمي إلى المجال  $[a, b]$ :

$$P(a < x \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$



حساب الاحتمال من خلال دالة التوزيع

➤ مثال:

أوجد قيمة الثابت  $C$  التي تحقق الشرطين الأول والثاني لدالة الكثافة الاحتمالية في الدالة التالية:

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & \text{دون ذلك} \end{cases}$$

- ✓ أحسب احتمال أن تكون  $X$  تنتمي للمجال من 1 إلى 2.
- ✓ أحسب احتمال أن تكون  $X$  لا تنتمي للمجال من 1 إلى 2.
- ✓ استخدم دالة التوزيع لحساب الاحتمال:  $P(1 < x < 2)$ .

الحل:

- إيجاد قيمة الثابت  $C$

$$\int_{-\infty}^{+\infty} f(x) dx = 1 \Rightarrow \int_{-\infty}^0 0 dx + \int_0^3 Cx^2 dx + \int_3^{+\infty} 0 dx = 1 \Rightarrow C \left[ \frac{x^3}{3} \right]_0^3 = 9C = 1 \Rightarrow C = 1/9$$

حتى تكون  $f(x)$  دالة كثافة يجب أن يكون  $C = 1/9$ .

➤ احتمال أن تكون  $x$  تنتمي للمجال من 1 إلى 2.



$$P(1 < x \leq 2) = \int_1^2 f(x)dx = \int_1^2 (1/9)x^2 dx = \frac{1}{9} \left[ \frac{x^3}{3} \right]_1^2 = \frac{1}{9} \left[ \frac{8-1}{3} \right] = \frac{7}{27}$$

➤ احتمال أن تكون  $x$  لا تنتمي للمجال من 1 إلى 2

$$P(1 > x > 2) = 1 - P(1 < x < 2) = 1 - \frac{7}{27} = \frac{20}{27}$$

➤ دالة التوزيع:

$$* x < 0: \quad F(x) = \int_{-\infty}^x f(u)du = \int_{-\infty}^0 0du = 0$$

$$* x \geq 3: \quad F(x) = \int_{-\infty}^x f(u)du = \int_{-\infty}^0 0du + \int_0^3 \frac{1}{9}u^2 du + \int_3^x 0du = 0 + \frac{1}{9} \left[ \frac{u^3}{3} \right]_0^3 + 0 = \frac{27}{27} = 1$$

$$* 0 \leq x < 3: \quad F(x) = \int_{-\infty}^x f(u)du = \int_{-\infty}^0 0du + \int_0^x \frac{1}{9}u^2 du = \frac{1}{9} \left[ \frac{u^3}{3} \right]_0^x = \frac{x^3}{27}$$

$$F(x) = \begin{cases} 0 & x < 0 \\ x^3 / 27 & 0 \leq x < 3 \\ 1 & x \geq 3 \end{cases} \quad \text{فتكون الصيغة النهائية لدالة التوزيع:}$$

➤ حساب الاحتمال:  $P(1 < x < 2)$  باستعمال دالة التوزيع:

$$P(1 < x \leq 2) = F(2) - F(1) = \frac{2^3}{27} - \frac{1^3}{27} = \frac{7}{27}$$

### 3.1- التوقع الرياضي:

إن التوقع الرياضي (الأمل الرياضي) لأي توزيع احتمالي معين هو عبارة عن متوسط حسابي لهذا التوزيع، و يرمز له بالرمز:  $E(X)$ .

أ) - حالة التوزيع الإحتمالي المنفصل:

ويعطى بالعلاقة التالية:

$$E(X) = \frac{\sum Xp(X)}{\sum P(X)}$$

و حيث أن:  $\sum P(X) = 1$  ، فإن:

$$E(X) = \sum XP(X) \quad (1.17)$$

➤ مثال:

تنتج شركة معينة 20 قطعة في اليوم من منتج معين، من بينهما قطعتان فاسدتان. قمنا بسحب 4 قطع من هذا المنتج بطريقة عشوائية. أوجد القيمة المتوقعة لعدد القطع الفاسدة.

➤ الحل:

إن قيم  $X$  الممكنة هي:  $X = \{0,1,2\}$  ، أي لا توجد قطعة فاسدة أو توجد قطعة فاسدة واحدة أو توجد قطعتان فاسدتان.

إن عدد الحالات الكلية هو:  $C_{20}^4 = 4845$  ؛ أما الإحتمالات المقابلة لـ  $X$  فهي:

$$P(X = 0) = \frac{C_2^0 C_{18}^4}{C_{20}^4} = 0.63$$

$$P(X = 1) = \frac{C_2^1 C_{18}^3}{C_{20}^4} = 0.33$$

$$P(X = 2) = \frac{C_2^2 C_{18}^2}{C_{20}^4} = 0.031$$

نتحصل على التوزيع التالي:

المجموع	2	1	0	قيم X
1	0.031	0.33	0.63	P (X)
0.392 = 0.4	0.062	0.33	0	XP(X)

وعليه فإن التوقع الرياضي لهذه العملية هي 0.4، أي أننا نتوقع الحصول على 0.4 قطعة فاسدة (قيمة لـ X) من بين 20 قطعة المنتجة.

(ب) - حالة التوزيع الإحتمالي المتصل:

تعطى عبارة التوقع الرياضي بالعلاقة التالية:

$$E(x) = \int_{\mathbb{R}} xf(x)dx \quad (1.18)$$

➤ مثال

يمكننا بسهولة إيجاد الأمل الرياضي للتوزيع الإحتمالي المتصل الوارد في المثال السابق وهو:

$$f(x) = \begin{cases} 1/4x & 1 \leq x \leq 3 \\ 0 & \text{دون ذلك} \end{cases}$$

➤ الحل:

$$\begin{aligned} E(x) &= \int_{\mathbb{R}} xf(x)dx = \int_1^3 \frac{1}{4}x^2dx = \frac{1}{4} \int_1^3 x^2dx = \frac{1}{4} \left[ \frac{x^3}{3} \right]_1^3 \\ &= \frac{1}{4} \left[ \frac{27}{3} - \frac{1}{3} \right] = \frac{26}{12} = 2.16 \end{aligned}$$

### ج) - خواص التوقع الرياضي:

- إذا كان  $X$  متغيرا عشوائيا غير سالب فإن:  $E(X) \geq 0$
- $E(C) = C$ ، حيث  $C \in R$
- $E(C.X) = C E(X)$
- $E(X.Y) = E(X).E(Y)$ ، إذا كانت المتغيران  $X$  و  $Y$  مستقلتان
- إذا كان  $X$  و  $Y$  متغيران عشوائيان معرفين على نفس فضاء العينة، فإن:  $E(X + Y) = E(X) + E(Y)$ .

### ➤ مثال:

يمكن لأحدى الشركات التجارية أن تربح 300.000 دولار باحتمال قدره 0.60، كما يمكنها أن تخسر 100.000 دولار باحتمال قدره 0.40. المطلوب: كم تتوقع الشركة أن تربح في المتوسط؟.

### ➤ الحل:

نلاحظ وجود متغيرين هما الربح  $X$  و الخسارة  $Y$ ، و لكل منهما احتمال معلوم. إن التوقع المشتركة لهذه العملية هي:

$$E(X + Y) = E(X) + E(Y) = 300.000 \times 0.60 + (-100.000 \times 0.40) = 140.000 \$$$

أي أن الشركة تتوقع أن تربح 140.000 دولار (لاحظ إشارة السالب التي تعني الخسارة).

### 4.1- التباين:

#### أ) - حالة التوزيع الإحتمالي المنفصل:

ويعطى بالعلاقة التالية:

$$\text{Var}(X) = \sum (X - E(X))^2.P(X) \quad (1.19)$$

وبعبارة أخرى، يمكن أن نكتب:

$$\text{Var}(X) = \sum x^2.P(x) - [E(x)]^2$$

(ب) - حالة التوزيع الإحتمالي المتصل:

$$\text{Var}(X) = \int_{\mathbb{R}} (X - E(X))^2 f(x) dx \quad (1.20)$$

أو نكتب:

$$\text{Var}(X) = \int_{\mathbb{R}} x^2 \cdot P(x) \cdot dx - [E(x)]^2$$

➤ تمرين:

يصوب رامي على هدف 03 طلقات، فإذا كان إحتمال إصابة الهدف هو 0.40. نعرف  $X$  المتغير العشوائي الذي يمثل عدد إصابات الهدف. المطلوب إيجاد التوقع الرياضي و التباين لهذه التجربة الإحتمالية.

➤ الحل:

إن قيم  $X$  الممكنة هي:

$$X = \{0, 1, 2, 3\}$$

أما الإحتمالات الممكنة فهي كما يلي:

$$\begin{aligned} P(X = 0) &= 0.6 \times 0.6 \times 0.6 = 0.216 \\ P(X = 1) &= 0.6 \times 0.6 \times 0.4 = 0.432 \\ P(X = 2) &= 0.6 \times 0.4 \times 0.4 = 0.288 \\ P(X = 3) &= 0.4 \times 0.4 \times 0.4 = 0.064 \end{aligned}$$

نلخص هذا في الجدول التالي:

المجموع	3	2	1	0	قيم $X$
1	0.064	0.288	0.432	0.216	$P(X)$

التوقع الرياضي:

$$E(X) = \sum X P(X) = 0 \times 0.216 + 1 \times 0.432 + 2 \times 0.288 + 3 \times 0.064 = 1.2$$

التباين:

$$Var(X) = \sum (X - E(X))^2 \cdot P(X) = [ (0 - 1.2)^2 \times 0.216 + (1 - 1.2)^2 \times 0.432 + (2 - 1.2)^2 \times 0.288 + (3 - 1.2)^2 \times 0.064 ] = 0.72$$

### 5.1- التوزيعات الإحتمالية الشهيرة:

هناك العديد من التوزيعات الاحتمالية التي تخضع إلى قوانين معينة ثابتة، نسبت في الغالب إلى واضعيها كتوزيع ذو الحدين الذي ينسب إلى العالم برنولي؛ ولكل هذه التوزيعات إستعمالاتها خاصة عند إجراء التجارب العلمية، إذ تستعمل في اختبار نجاح هذه التجارب، كما تساعد في إتخاذ القرار. نكتفي بدراسة بعض التوزيعات الشهيرة المستعملة أساسا في المجال الاقتصادي وكذلك تلك التي تستعمل في تجارب العلوم الطبيعية، وهي:

#### 1.5.1- توزيع ذو الحدين:

وهو أحد التوزيعات الإحتمالية المنفصلة؛ فإذا كان لدينا في تجربة إحتمالية مستقلة ناتجين، نفرض أن  $P$  هي إحتمالات النجاح، فيكون  $q$  هي إحتمالات الفشل بحيث:  $p + q = 1$ . إن إحتمال أن وقوع  $X$  من النجاحات في  $n$  من المحاولات المتكررة هو:

$$P(X = x) = C_n^x p^x q^{n-x} = \frac{n!}{(n-x)!x!} p^x q^{n-x} \quad (1.21)$$

ونرمز لهذا الإقتران الإحتمالي بالرمز:  $b(x, n, p)$   
➤ مثال:

ترمى قطعة نقود 6 مرات. أوجد إحتمال الحصول على:

- الوجه  $F$  مرتين بالضبط
- الحصول 4 مرات على الوجه  $F$  على الأقل
- عدم الحصول على الوجه  $F$ .



➤ **الحل:**

لدينا في قطعة النقود:  $p = q = 0.5$ ، فإذا كانت  $n = 6$  فإن:

- احتمال الحصول على الوجه F مرتين بالضبط:

$$P(X = x) = C_n^x p^x q^{n-x} \Rightarrow P(X = 2) = C_6^2 0.5^2 0.5^{6-2} = 0.234$$

- الحصول 4 مرات على الوجه F على الأقل:

$$P(X = 4) + P(X = 5) + P(X = 6)$$

$$= C_6^4 \cdot 0.5^4 \cdot 0.5^{6-4} + C_6^5 \cdot 0.5^5 \cdot 0.5^{6-5} + C_6^6 \cdot 0.5^6 \cdot 0.5^{6-6} = 0.343$$

- عدم الحصول على الوجه F.

$$P(X = 0) = C_6^0 0.5^0 0.5^6 = 0.5^6 = 0.0156$$

1.1.5.1- خواص توزيع ذو الحدين:

• التوقع الرياضي:

$$\mu = np \quad (1.22)$$

• التباين:

$$\sigma^2 = npq \quad (1.23)$$

• الانحراف المعياري:

$$\sigma = \sqrt{npq} \quad (1.24)$$

➤ **تمرين:**

ألقي حجر نرد 180 مرة. ما هو توقع عدد مرات الحصول على الرقم 6 و ما هو الانحراف المعياري عندئذ؟.

➤ **الحل:**

$$\mu = np = 180 \times 1/6 = 30$$

أي أننا نتوقع 30 مرة الحصول على الرقم 6، بانحراف معياري:

$$\sigma = \sqrt{npq} = \sqrt{180 \times 1/6 \times 5/6} = 5$$

2.5.1- توزيع بواسون:

وهو أحد التوزيعات الاحتمالية المنفصلة، وهو توزيع لا نهائي قابل للعد، له إستعمالات في الظواهر المتعلقة بالزمن، مثل حساب احتمالات عدد

السيارات التي تمر خلال دقيقة في منطقة ما؛ أو حساب عدد الجسيمات أشعة ألفا التي يطلقها مركب مشع خلال وحدة زمنية...

ولإيجاد احتمالات المتغير العشوائي  $X$  المقابلة لهذا التوزيع، فإننا نستخدم العلاقة التالية:

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad (1.25), \quad x = 0, 1, 2, \dots$$

حيث أن  $\lambda$  ثابت أكبر من الصفر يمثل متوسط عدد الأحداث؛  $e$  أساس اللوغارتم النيفري و يساوي 2.71.

#### 1.2.5.1- خواص توزيع بواسون:

• التوقع الرياضي:

$$\mu = \lambda \quad (1.26)$$

• التباين:

$$\sigma^2 = \lambda \quad (1.27)$$

• الانحراف المعياري:

$$\sigma = \sqrt{\lambda} \quad (1.28)$$

➤ مثال:

يتلقى قسم الشرطة متوسط 5 مكالمات في الساعة. أوجد احتمال تلقي مكالمتين في ساعة مختارة عشوائيا.

➤ الحل:

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{5^2 e^{-5}}{2!} = 0.08$$

#### 2.2.5.1- تقريب توزيع ذو الحدين من توزيع بواسون:

في كثير من الحالات تكون قيمة  $n$  كبيرة جدا في توزيع ذو الحدين، وتكون قيمة  $p$  (إحتمالات النجاح) صغيرة جدا؛ فيصبح صعبا إجراء الحسابات

(مثلا حساب  $n!$ ). يمكننا إجراء تقريب ذو الحدين لتوزيع بواسون بحيث أن:  $\lambda = np$ .

### ➤ مثال:

لنفترض أن  $n = 100$  و  $p = 0.01$ ، فيكون  $\lambda = np = 1$ . نتحصل على الجدول المقارن التالي:

x	0	1	2	3	4	5
إحتمالات ذو الحدين	0.366	0.370	0.185	0.0610	0.0149	0.0029
إحتمالات بواسون	0.368	0.368	0.184	0.0613	0.0153	0.00307

يظهر جليا بعد مقارنة إحتمالات التوزيعين، مدى التقارب الشديد بينهما.

### ➤ مثال 2:

يحتوي كتاب على 500 صفحة، بفرض أن هناك 300 خطأ مطبعيا، فأوجد إحتمال أن تحتوي صفحة معينة على:  
(أ) - خطأين بالضبط، (ب) - خطأين أو أكثر

### ➤ الحل:

إن هذه التجربة الإحتمالية تتبع توزيع ذو الحدين مع:  $p = n = 300$   
 $1/500$  وحيث أن قيمة  $n$  كبيرة و  $p$  صغيرة، فإننا نستخدم توزيع بواسون كتقريب لذو الحدين مع  $\lambda = np = 0.6$ .

(أ) - إحتمال الحصول على خطأين بالضبط:

$$P(X = 2) = \frac{0,6^2 e^{-0,6}}{2!} = 0.098$$

ب) - احتمال الحصول على خطأين أو أكثر: يعني صفر خطأ أو خطأ واحد أو خطأين:

$$P(X=0) + P(X=1) + P(X=2) = \frac{0,6^0 e^{-0,6}}{0!} + \frac{0,6^1 e^{-0,6}}{1!} + \frac{0,6^2 e^{-0,6}}{2!}$$

$$= 0.549 + 0.329 + 0.098 = 0.976$$

### 3.5.1- التوزيع الطبيعي:

أحد و أهم التوزيعات الإحصائية المتصلة، نظرا لإستعمالاته المتعددة في شتى الميادين. تعطى دالة كثافته الإحصائية بالعلاقة التالية:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2/\sigma^2} \quad (1.29)$$

حيث:  $\mu$  و  $\sigma > 0$  هما ثابتان. إن منحنى هذه الدالة يأخذ الشكل الجرسى، متمائل حول نقطة مركزية هي  $\mu$ ، بحيث أن المساحة تحت المنحنى تساوي 1. نرسم للتوزيع الطبيعي الذي وسطه  $\mu$  و تباينه  $\sigma^2$  بـ  $N(\mu, \sigma^2)$ .

#### 1.3.5.1- خواص التوزيع الطبيعي:

- التوقع الرياضي:  $\mu$
- التباين:  $\sigma^2$
- الانحراف المعياري:  $\sigma$

#### 2.3.5.1- التوزيع الطبيعي القياسي:

إن حساب احتمال  $P(x_1 < X < x_2)$  يتطلب حساب التكامل المحدود للدالة المعطاة في (1.29). إن حساب هذا التكامل يتطلب الكثير من الحسابات؛ ولذلك ظهر التوزيع الطبيعي القياسي الذي يعتمد في حساب هذا التكامل على إستعمال جدول خاص يمكن من حساب أي مساحة من خلال قراءة مباشرة في الجدول دون اللجوء إلى حساب هذا التكامل. ويمكن الانتقال من التوزيع

الطبيعي  $N(\mu, \sigma^2)$  إلى التوزيع الطبيعي القياسي الذي وسطه 0 و تباينه 1 أي  $N(0,1)$  بالعلاقة التالية :

$$Z = \frac{X - \mu}{\sigma} \quad (1.30)$$

عندئذ تصبح الدالة  $f(x)$  كما يلي:

$$\Psi(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2} \quad (1.31)$$

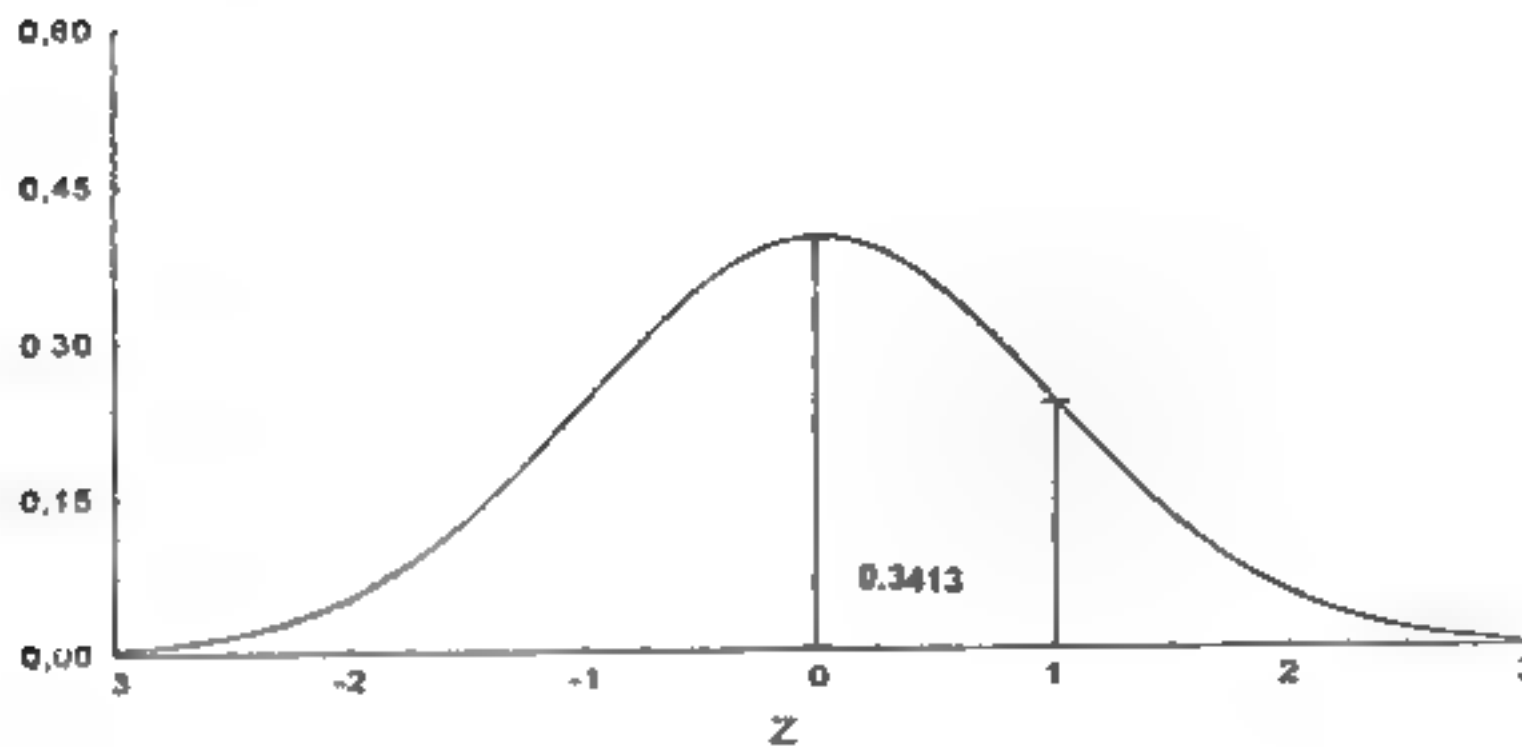
يعطينا جدول التوزيع الطبيعي القياسي (أنظر الملحق، جدول 3) المساحة الواقعة تحت المنحنى بين  $Z = 0$  وأي قيمة موجبة أخرى  $Z$ ، كما يمكننا الاستفادة من خاصية التماثل (التناظر) في هذا التوزيع من تحويل وحدات  $Z$  السالبة إلى وحدات موجبة.

➤ مثال 1:

-أحسب احتمال  $P(0 < Z < 1)$ :

إن هذا الاحتمال يقابل المساحة المحصورة بين 0 و 1.00 في الجدول 3. نأخذ العمود من أعلى إلى أسفل و نتوقف عند القيمة 1.0، الرقم الثاني وراء الفاصل يقرأ في سطر الجدول عند 0 فنجد المساحة: 0.3413 عند إلتقاء القراءتين.

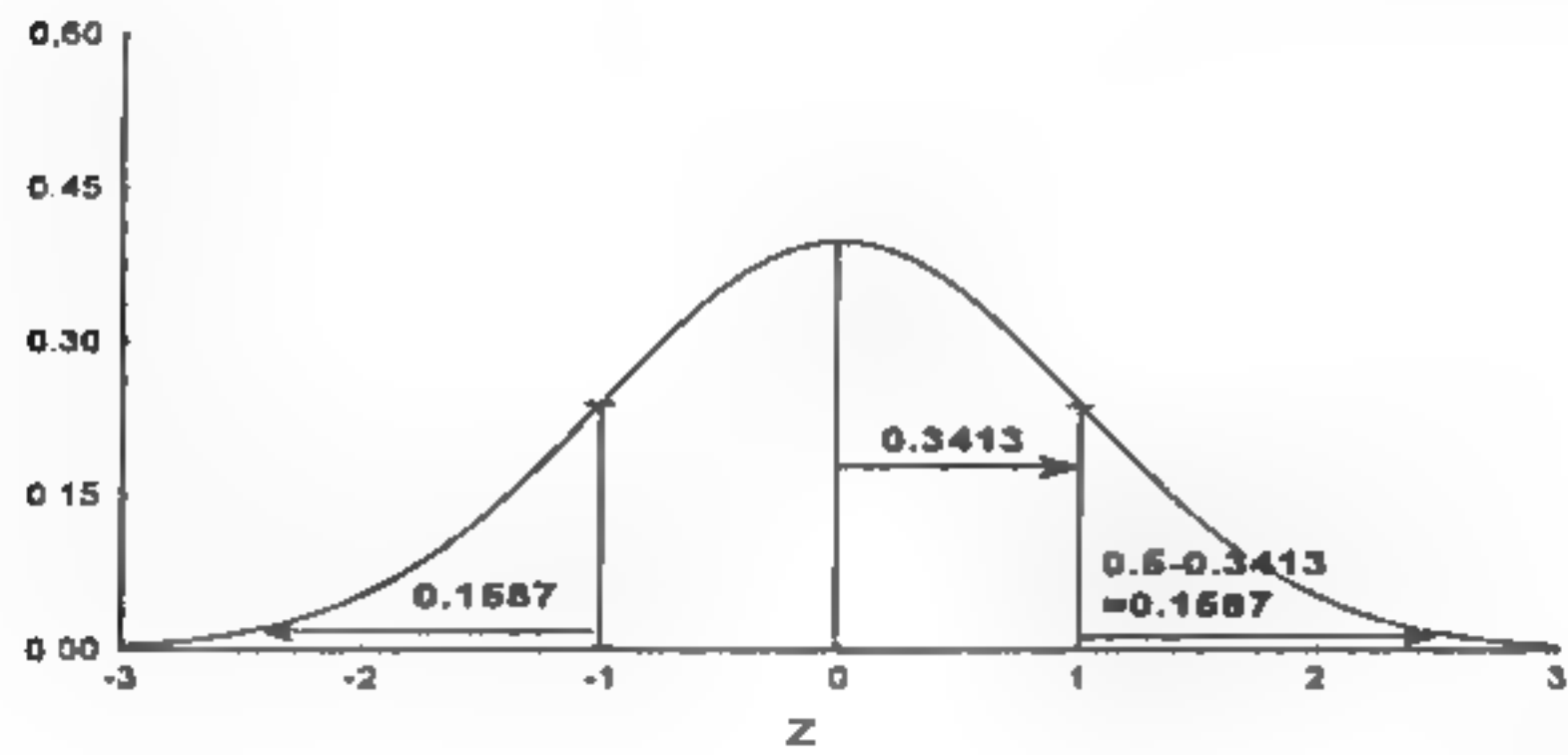
$$P(0 \leq Z \leq 1) = 0.3413$$



مثال 2:

أحسب احتمال  $P(Z \geq 1)$

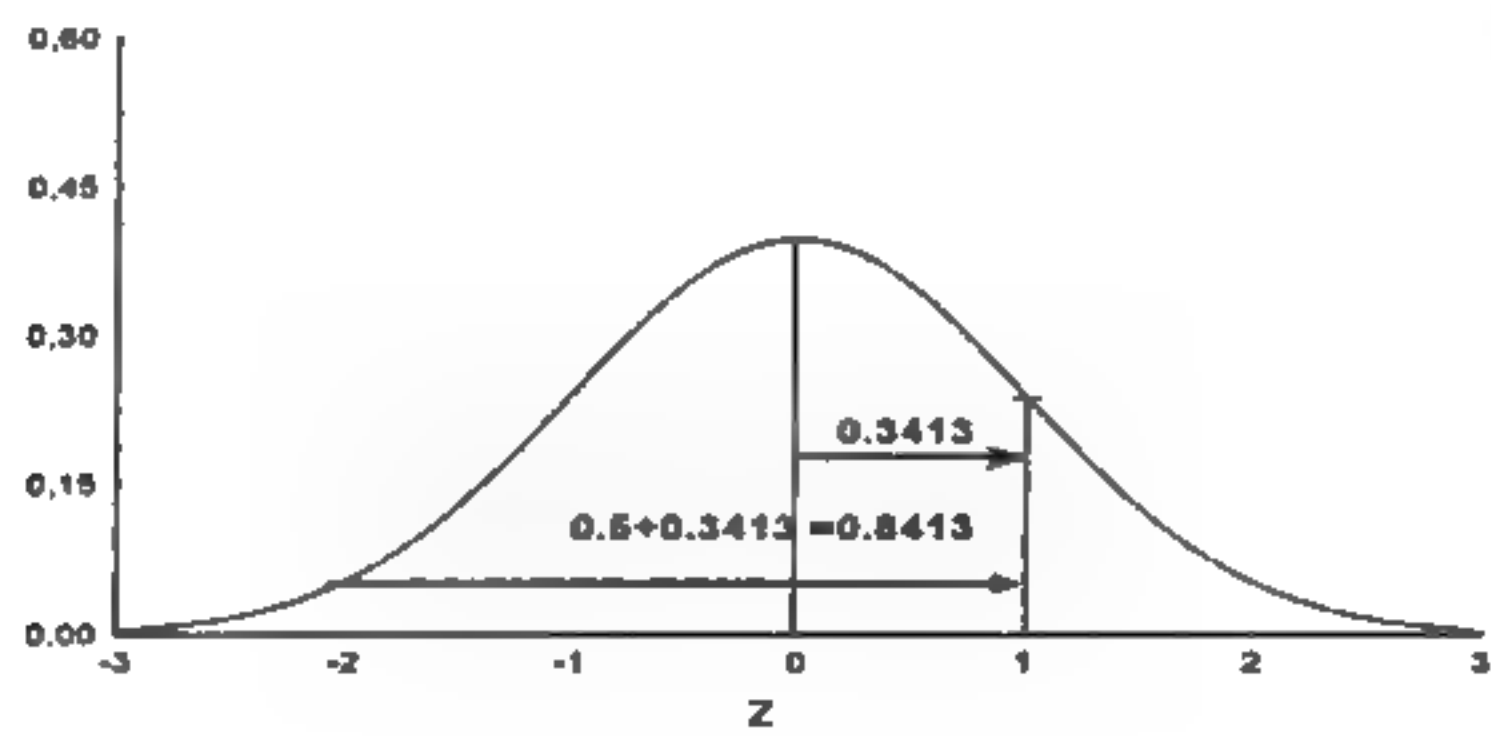
$$P(Z \geq 1) = 0.5 - (0 \leq Z \leq 1) = 0.5 - 0.3413 = 0.1587$$



➤ مثال 3:

أحسب احتمال  $P(Z \leq 1)$

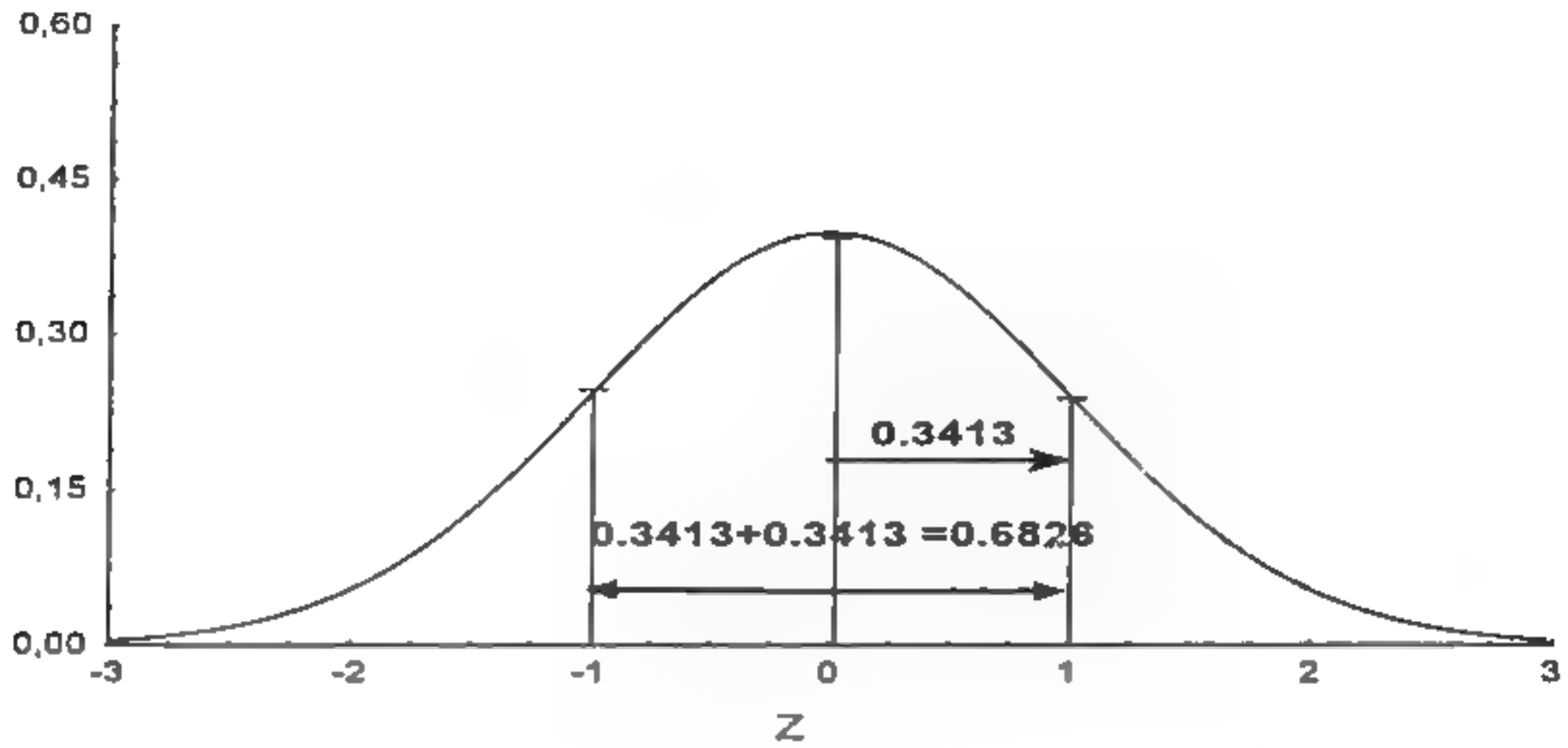
$$P(Z \leq 1) = 0.5 + P(0 \leq Z \leq 1) = 0.5 + 0.3413 = 0.8413$$



مثال 4:

أحسب احتمال  $P(-1 \leq Z \leq 1)$

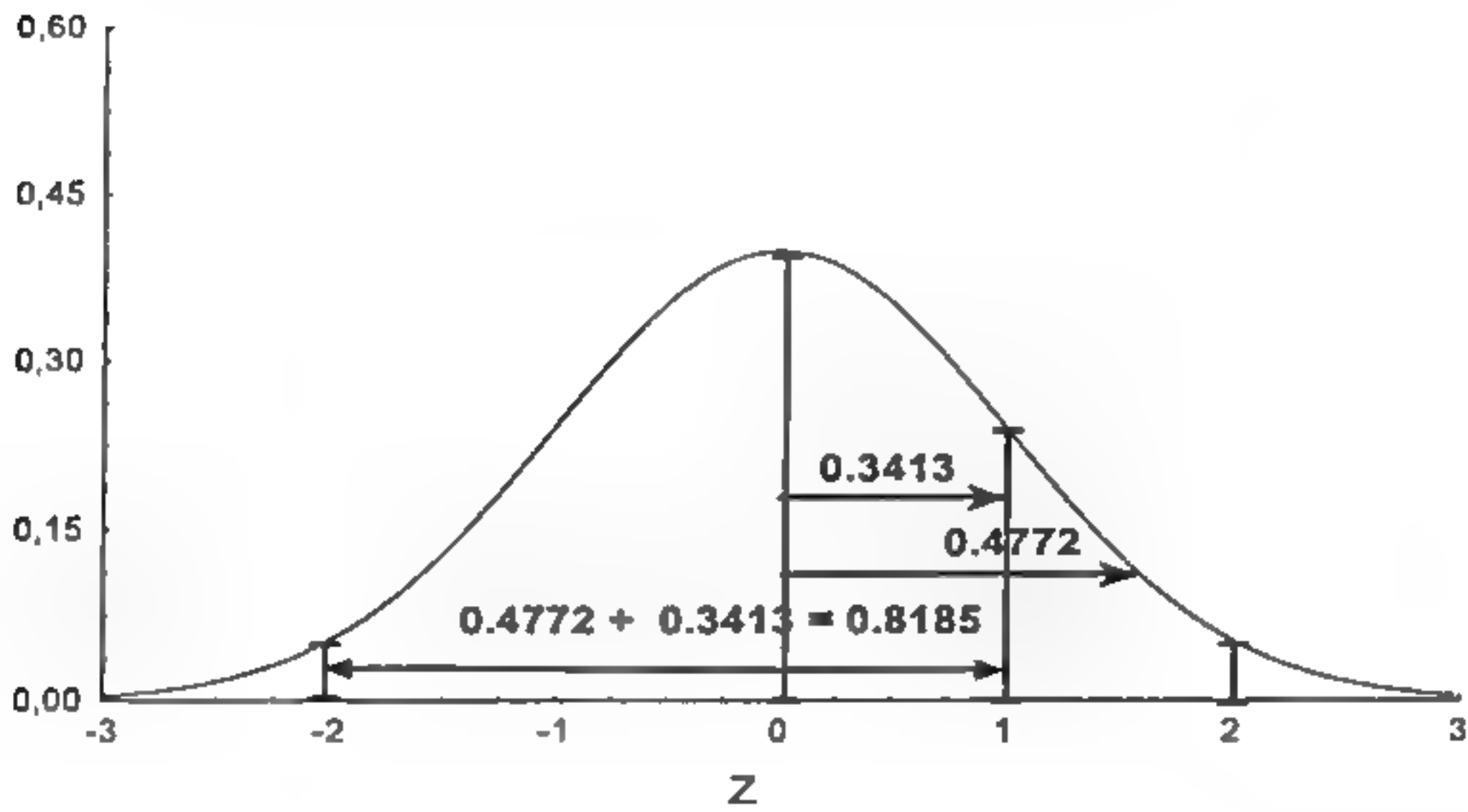
$$P(-1 \leq Z \leq 1) = 2 P(0 \leq Z \leq 1) = 2 \times 0.3413 = 0.6826$$



مثال 5:

أحسب احتمال  $P(-2 \leq Z \leq 1)$ :

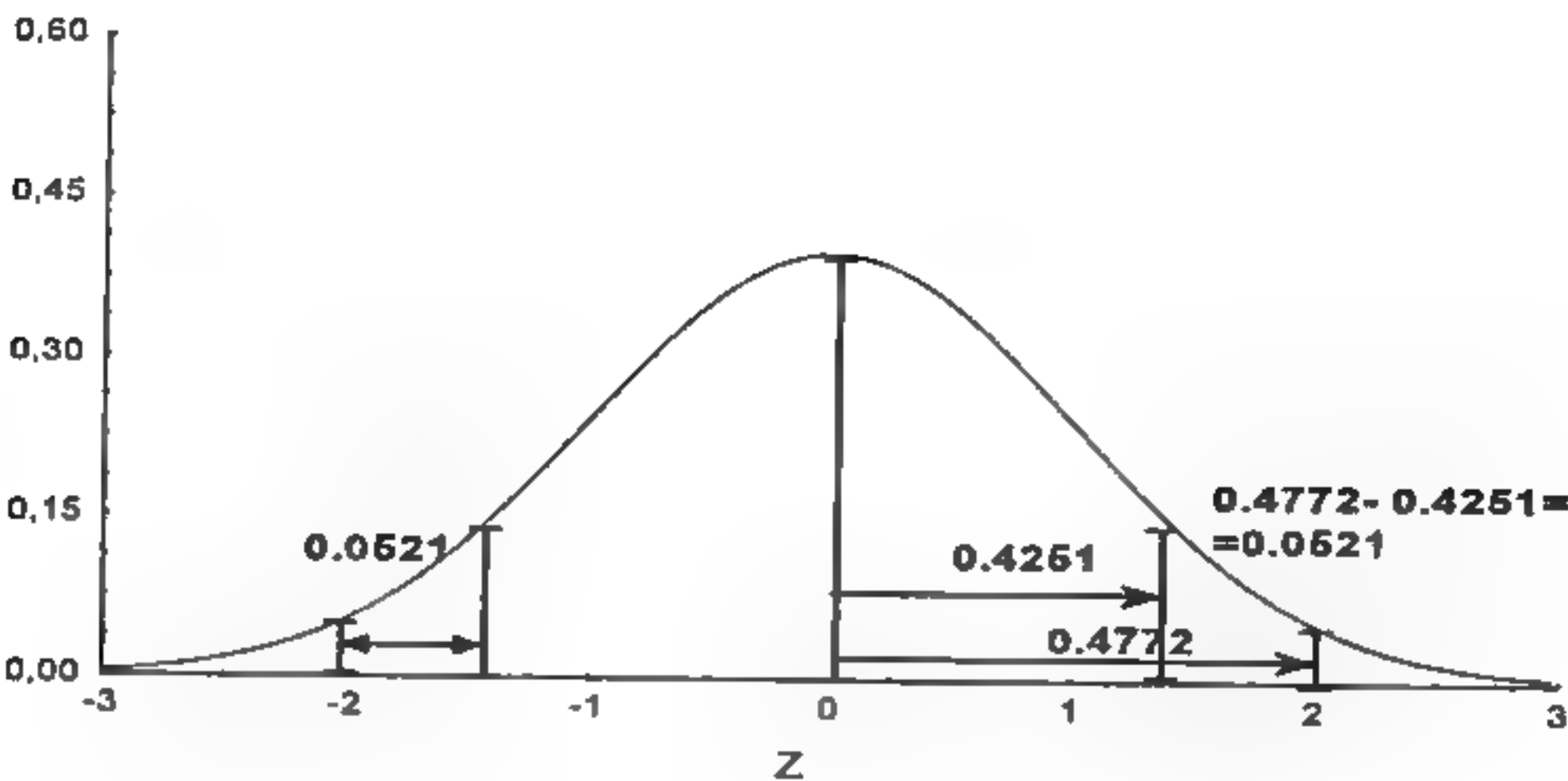
$$\begin{aligned} P(-2 \leq Z \leq 1) &= P(0 \leq Z \leq 2) + P(0 \leq Z \leq 1) = \\ &= 0.4772 + 0.3413 = 0.8185 \end{aligned}$$



مثال 6:

أحسب احتمال  $P(-2 \leq Z \leq -1.44)$ :

$$\begin{aligned} P(-2 \leq Z \leq -1.44) &= P(0 \leq Z \leq 2) - P(0 \leq Z \leq 1.44) = \\ &= 0.4772 - 0.4251 = 0.0521 \end{aligned}$$

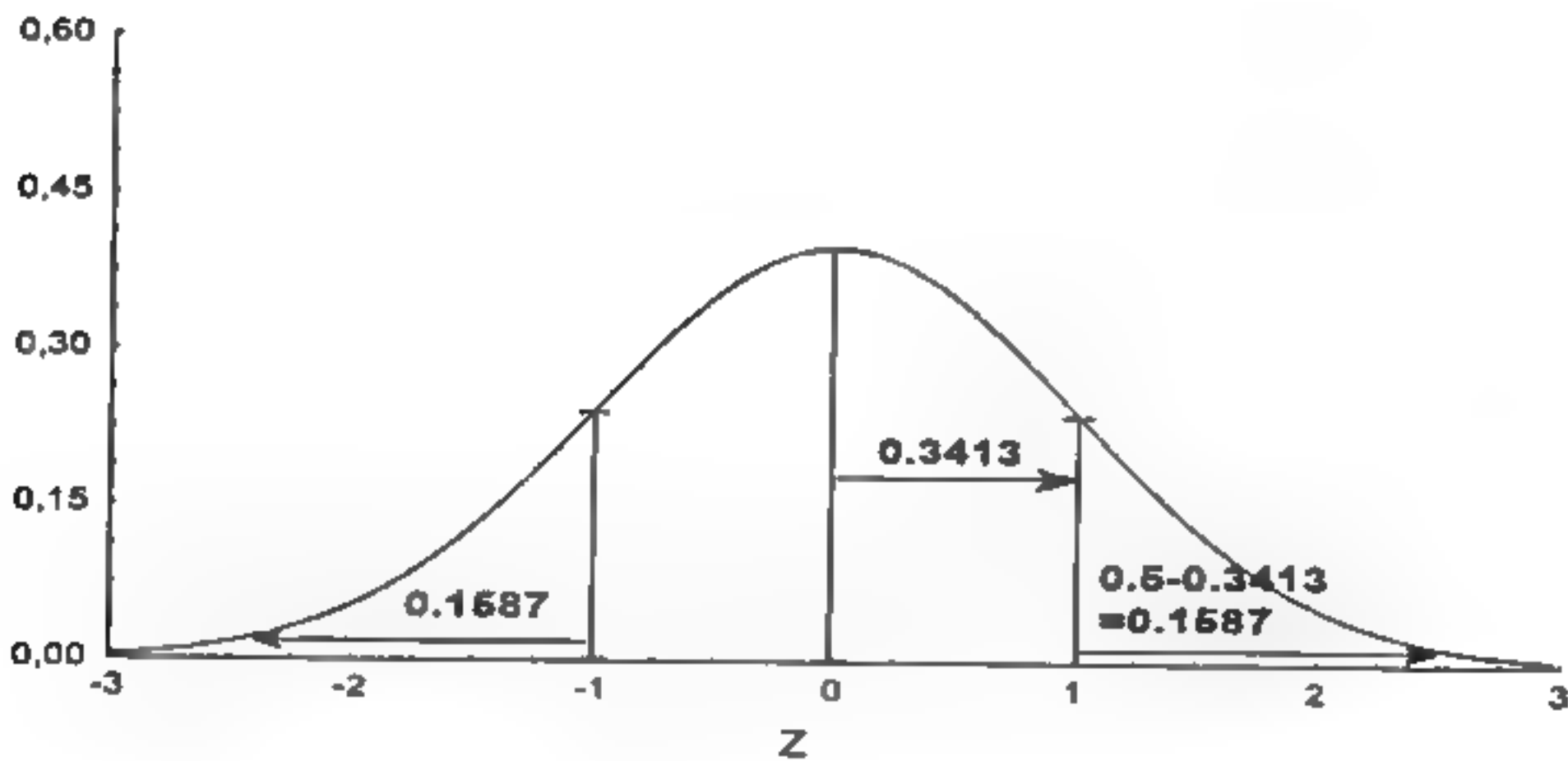


➤ مثال 7:

أحسب احتمال  $P(Z \leq -1)$

$$P(Z \leq -1) = P(Z \geq 1) = 0.5 - P(0 \leq Z \leq 1)$$

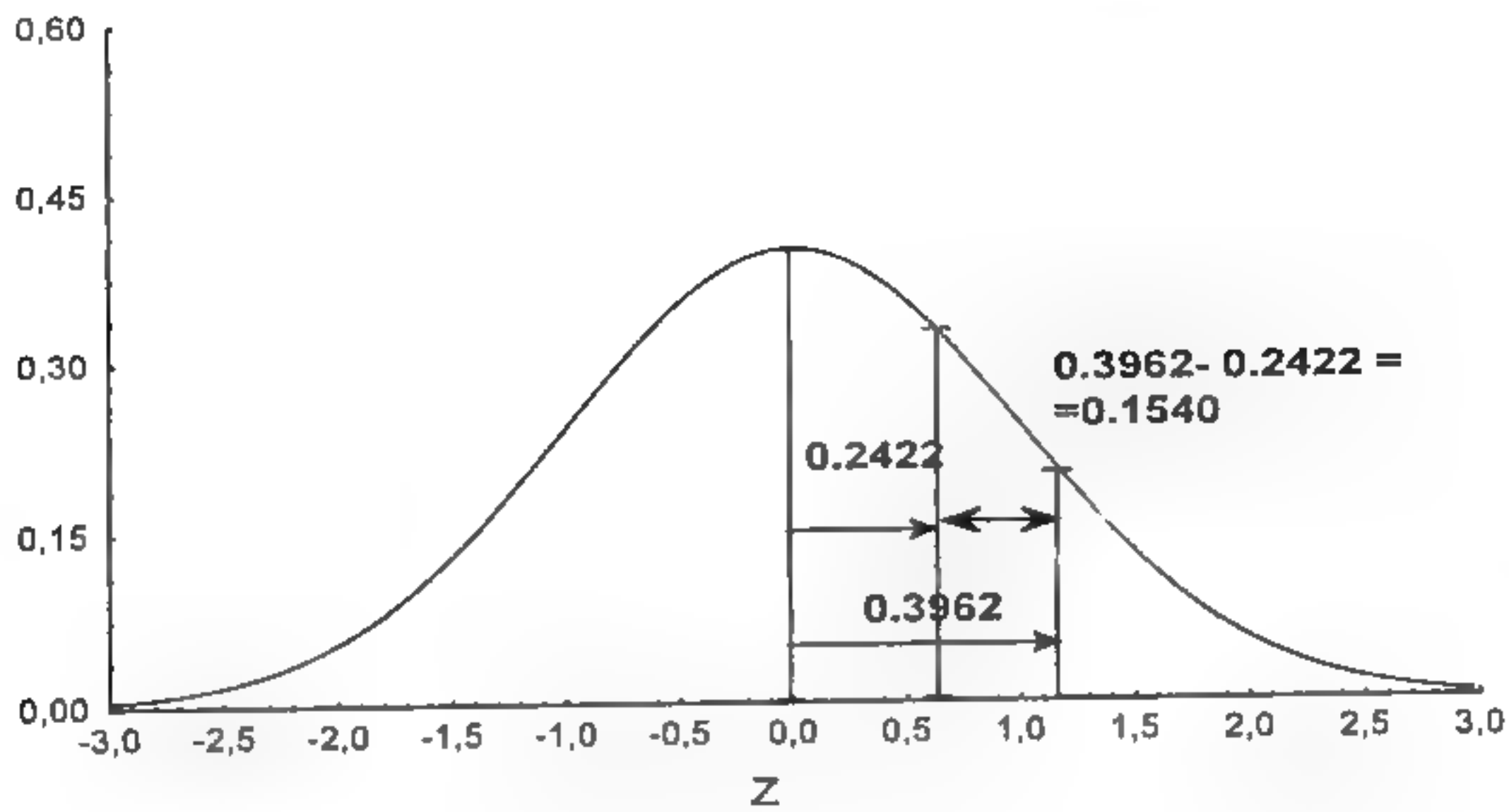
$$= 0.5 - 0.3413 = 0.1587$$



➤ مثال 8:

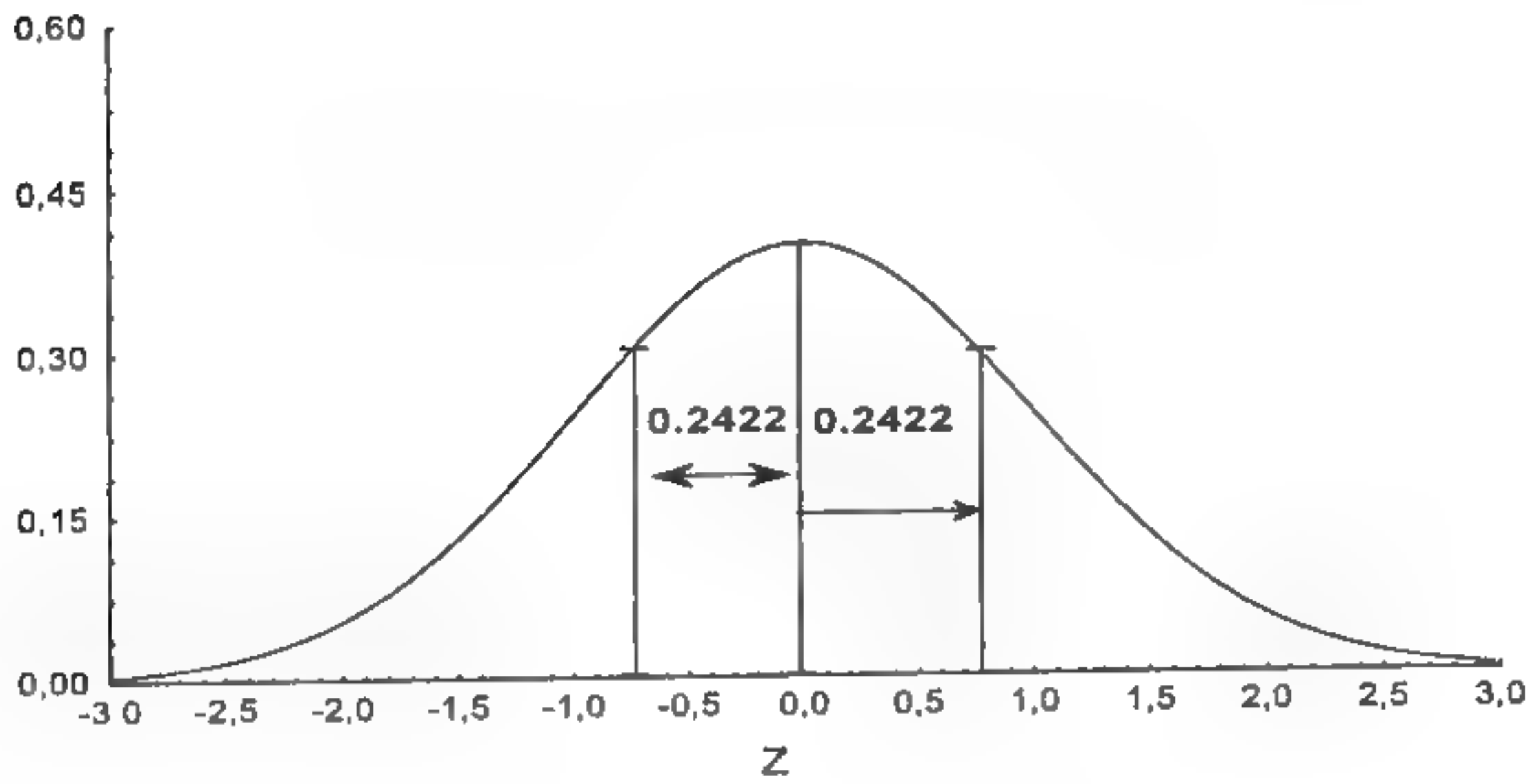
$$\begin{aligned} P(0.65 \leq Z \leq 1.26) &= P(0 \leq Z \leq 1.26) - P(0 \leq Z \leq 0.65) = \\ &= 0.3962 - 0.2422 = 0.1540 \end{aligned}$$





➤ مثال 9:

$$P(-0.65 \leq Z \leq 0) = P(0 \leq Z \leq 0.65) = 0.2422$$



➤ مثال 10:

إذا كانت درجة الحرارة تخضع للتوزيع الطبيعي بتوقع 20 دم و إنحراف معياري 3.33 دم. أوجد احتمال أن تتراوح درجة الحرارة بين 21.11 دم و 26.66 دم.

➤ الحل:

في هذه الحالة يجب تحويل وحدات التوزيع الطبيعي  $N(20, 3.33)$  إلى وحدات التوزيع الطبيعي القياسي  $N(0, 1)$  وفق العلاقة:

$$Z = \frac{X - \mu}{\sigma}$$

أي أن:

$$\begin{aligned} P(21.11 \leq X \leq 26.66) &= P\left(\frac{21.11 - 20}{3.33} \leq \frac{X - \mu}{\sigma} \leq \frac{26.66 - 20}{3.33}\right) = \\ P(0.33 \leq Z \leq 2) &= P(0 \leq Z \leq 2) - P(0 \leq Z \leq 0.33) \\ &= 0.4772 - 0.1293 = 0.3479 \end{aligned}$$

### 3.3.5.1- التقريب الطبيعي لتوزيع ذو الحدين:

إذا كان  $X$  متغيرا عشوائيا يخضع لتوزيع ذو الحدين  $(X, n, p)$ ، فإن احتمالات ذو الحدين تقترب من احتمالات التوزيع ذو الحدين، عندما تكون  $n$  كبيرة، حيث:

$$\frac{X - np}{\sqrt{npq}} \rightarrow N(0, 1) \quad (1.32)$$

وذلك حسب الخواص التالية:

- إذا كان  $k$  عددا صحيحا موجبا بين  $0$  و  $n$  فإن:

$$\begin{aligned} P(X=k) &= P(k - 1/2 \leq X \leq k + 1/2) = \\ \frac{k - 1/2 - np}{\sqrt{npq}} \leq Z &\leq \frac{k + 1/2 - np}{\sqrt{npq}} \quad (1.33) \end{aligned}$$

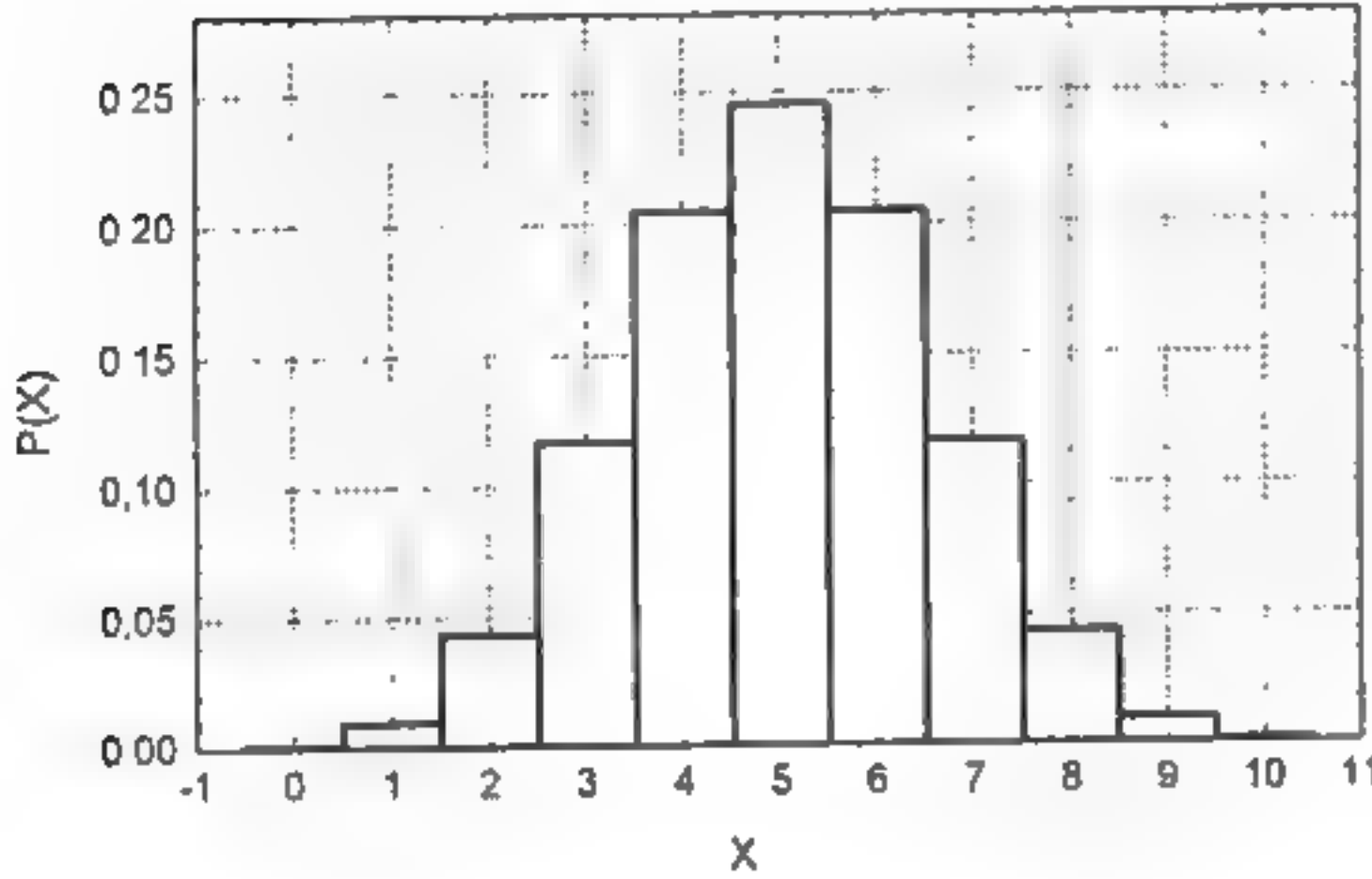
- إذا كان  $a$  و  $b$  أعدادا صحيحة موجبة بين  $0$  و  $n$  فإن:

$$(1.32)$$

$$P(a \leq X \leq b) = \frac{a - 1/2 - np}{\sqrt{npq}} \leq Z \leq \frac{b + 1/2 - np}{\sqrt{npq}}$$

## ➤ مثال:

نرسم المدرج الإحصائي لتوزيع ذو الحدين التالي:  $b(X, 10, 0.5)$ ، حيث:  $X=0,1,2,\dots,10$  مع:



$P(X=0) = 0.0010$
$P(X=1) = 0.0980$
$P(X=2) = 0.0439$
$P(X=3) = 0.1172$
$P(X=4) = 0.2051$
$P(X=5) = 0.2460$
$P(X=6) = 0.2050$
$P(X=7) = 0.1170$
$P(X=8) = 0.0430$
$P(X=9) = 0.0980$
$P(X=10) = 0.0011$

لنأخذ مثلا المستطيل رقم 5، نلاحظ أن عرضه يساوي  $5.5 - 4.5 = 1.0$ ، أما طوله فهو 0.246، وعليه تكون مساحته هي 0.246 (الطول في العرض)؛ وهكذا لباقي المستطيلات.

لنحسب هذه المساحات بالتوزيع الطبيعي وفق أحد القواعد السابقة، ولنأخذ مثلا المستطيل رقم 5.

لدينا:  $\mu = np = 10 \times 0.5 = 5$  و  $\sigma = \sqrt{npq} = \sqrt{10 \times 0.5 \times 0.5} = \sqrt{2.5}$  فيكون:

$$P(x = 5) = P\left(\frac{4.5 - 5}{\sqrt{2.5}} \leq Z \leq \frac{5.5 - 5}{\sqrt{2.5}}\right) = P(-0.316 \leq Z \leq 0.316) = 2 \times 0.1217 = 0.2434$$

من خلال المقارنة بين النتيجة، يظهر جليا إمكانية إجراء هذا التقارب.

### ➤ تمرين:

لنفرض أنه لدينا 300 سؤال، لكل سؤال 4 أجوبة، واحدة فقط صحيحة. أوجد احتمال أن يحصل الطالب على 78 إجابة صحيحة. أوجد احتمال أن يحصل الطالب على 72 و 100 إجابة صحيحة.

### ➤ الحل:

لدينا  $X$  عدد الإجابات الصحيحة،  $n=300$  و  $p = \frac{1}{4}$ .

$$\mu = np = 75$$

$$\sigma = 7.5$$

لنحسب هذا الاحتمال باستعمال ذو الحدين، فنحصل على:

$$b(78, 300, \frac{1}{4}) = C_{300}^{78} (\frac{1}{4})^{78} (\frac{3}{4})^{222} = \dots$$

إن حساب هذا الاحتمال يتطلب الكثير من الحسابات، و حيث أن  $n$  كبيرة فإنه يمكن تقريب ذو الحدين بالطبيعي:

$$P(x = 78) = P\left(\frac{77.5 - 75}{7.5} \leq Z \leq \frac{78.5 - 75}{7.5}\right) \\ = P(0.333 \leq Z \leq 0.467) = 0.051$$

ب)- احتمال أن يحصل الطالب على 72 و 100 إجابة صحيحة:

$$P(72 \leq x \leq 100) = P\left(\frac{71.5 - 75}{7.5} \leq Z \leq \frac{100.5 - 75}{7.5}\right) = 0.680$$

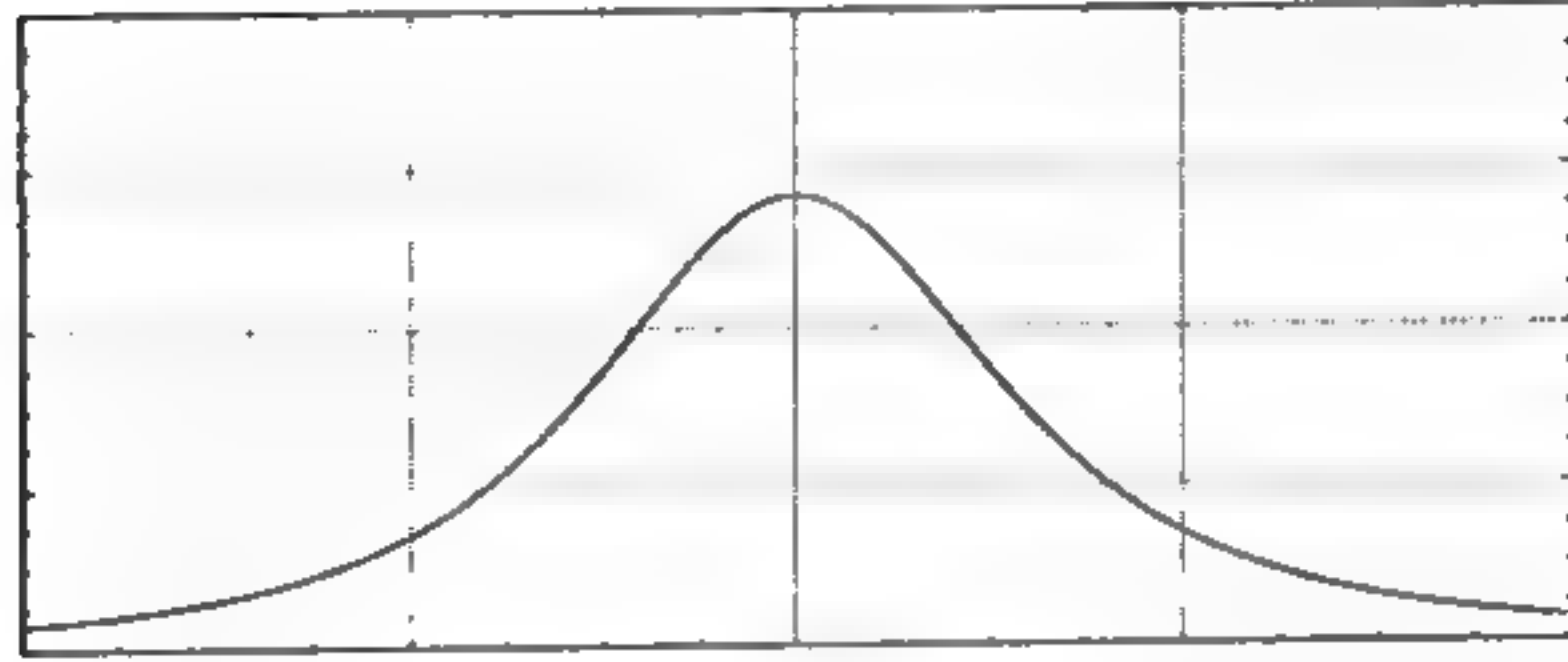
### 4.5.1- توزيع ستونت (t):

أحد التوزيعات الاحتمالية المتصلة، يشبه كثيرا التوزيع الطبيعي، تكمن تطبيقاته في نظرية العينات واختبار الفرضيات كبديل للتوزيع الطبيعي عندما يكون حجم العينة صغيرا.

دالة كثافته الاحتمالية تعطى بالعلاقة التالية:

$$f(t) = c(1 + \frac{t^2}{v})^{-(v+2)/2} \quad (1.34)$$

حيث:  $c$ : ثابت مرتبط بـ  $v$   
 $v$ : درجات الحرية و  $-\infty < t < +\infty$   
 نرمز لتوزيع  $t$  بالرمز  $t(p, v)$ .



تُحسب احتمالات توزيع  $t$  من خلال الجدول (أنظر الملحق، جدول 4)، حيث تسجل درجات الحرية في العمود الأيسر، أما في الأعلى فتقرأ المساحات (الاحتمالات) وفي داخل الجدول تقرأ قيم  $t$  المقابلة.

➤ مثال:

ما هي قيمة المساحة التي تقابل حجم العينة 11 و  $t = 2.812$ .

➤ الجواب:

لدينا  $v = n - 1$ ، أي أن  $v = 10$  و  $t = 2.812$  فإن المساحة المقابلة هي 0.95.

إن توزيع  $t$  هو توزيع متمائل، حيث أن:

$$\begin{aligned} \mu &= 0 \\ \sigma^2 &= \frac{v}{v-2} \end{aligned} \quad \text{حيث } v > 2 \quad (1.35)$$

\* درجات الحرية تعرف بأنها العدد  $n$  من المشاهدات المستقلة في العينة ناقص العدد  $k$  لمعالم المجتمع (الوسط الحسابي و الانحراف المعياري أي أن:  $v = n - k$  في حالة توزيع  $t$  فإن  $v = n - 1$  على أساس أنه يجب معرفة المتوسط الحسابي لهذا التوزيع والذي يساوي الصفر، و بالتالي يكون عدد المعالم المقدرة هو 1.

وبالتالي أمكن الاستفادة من هذه الخاصية في حساب المساحات الصغيرة  
(و الغير موجودة في الجدول أحيانا)، و ذلك بتطبيق العلاقة التالية:

$$t_{1-p} = -t_p \quad (1.36)$$

فمثلا، يمكن الكتابة:

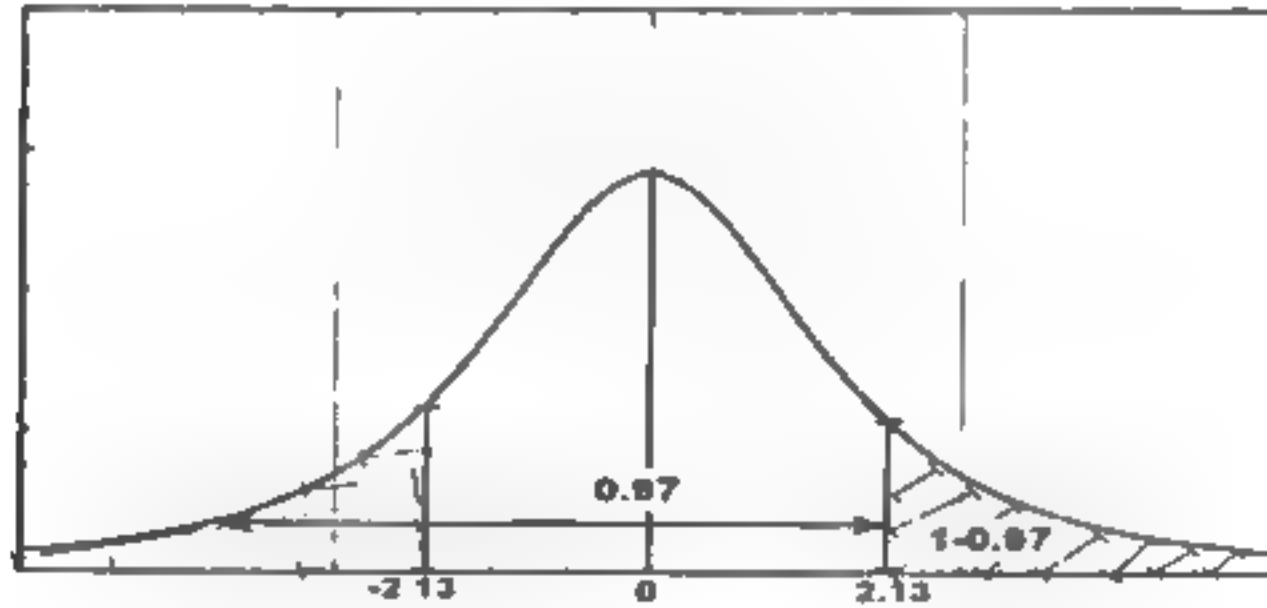
$$t_{0.05} = -t_{0.95}$$

➤ مثال:

بإستعمال توزيع ستودنت، أوجد المساحة الواقعة على يسار  $t = -2.13$   
وعند درجات حرية 15.  
لدينا:

$t_{1-p} = -t_p$ ، وحيث أن قيمة  $t$  سالبة فإننا نستخدم خاصية التناظر  
(لا توجد قيم سالبة في جدول  $t$ ) أي:

$$t(p,15) = -t(1-p,15)$$



إن قيمة  $t = -2.13$  تقابلها  $t = 2.13$  ومساحتها 0.97 وحيث أن  
المساحة المطلوبة هي على يسار -2.13، فيكون:

$1-p = 0.97$  من الجدول، ومنه  $p = 1-0.97 = 0.03$ ، أي أن المساحة  
الواقعة على يسار  $t = -2.13$  و عند درجات حرية 15 هي 0.03.

➤ مثال 2:

أوجد قيمة  $p$  بحيث يكون:  $t(p,15) = -2.60$ .

➤ الحل:

بالتماثل نجد:  $t(p,15) = -t(1-p,15)$ ، أي أن:  $-t(1-p,15) = -2.60$   
ومن الجدول:  $1-p = 0.99$  هذا يعني أن:  $p = 0.01$ .

➤ مثال 3:

أوجد قيمة  $t$  التي تقابل  $p = 0.05$  و  $v = 5$ .

➤ الحل:

حيث أن قيمة المساحة صغيرة جدا، يمكن تطبيق الخاصية التالية:

$$t(0.05,15) = -t(1-0.05,15)$$

$$t(0.05,15) = -t(0.95,15) = -2.015$$

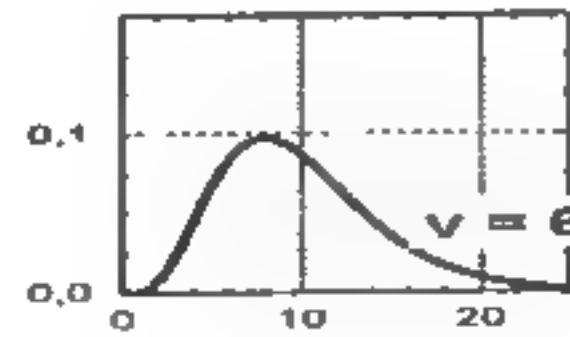
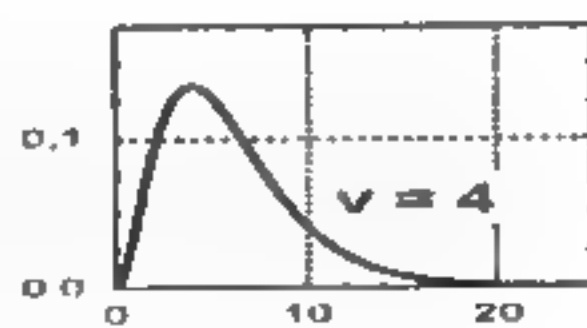
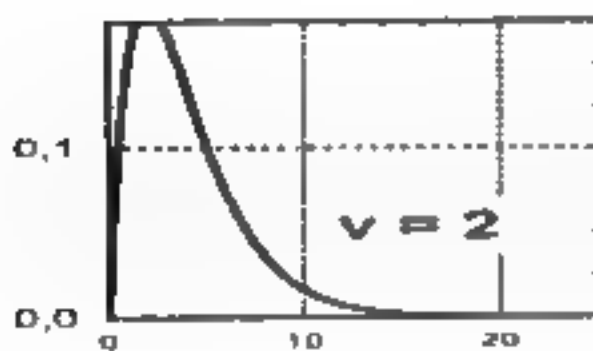
### 5.5.1- توزيع كاي مربع $\chi^2$ (khi deux) :

توزيع احتمالي متصل، له إستعمالات متعددة خاصة في إختبارات الارتباط و الإستقلال و التوفيق؛ و هو معرف بمتغيره العشوائي  $\chi^2$ . دالة كثافته الاحتمالية تعطى بالعلاقة التالية:

$$f(\chi^2) = c(\chi^2)^{(v-2)/2} e^{-\chi^2/2}, \quad \chi^2 > 0 \quad (1.37)$$

حيث أن  $c$  ثابت يعتمد على  $v$  ليجعل المساحة تحت المنحنى تساوي

$$1 \text{ و } v = n - 1.$$



ولإيجاد احتمالات الـ  $\chi^2$ ، فإننا نستخدم الجدول الخاص بهذا التوزيع (أنظر الملحق، جدول 5)، حيث نجد أفقيا المساحات المقابلة، عموديا درجات الحرية وفي داخل الجدول نقرأ قيم الـ  $\chi^2$ .

➤ مثال 1:  $(0.99, 10) = 23.20$

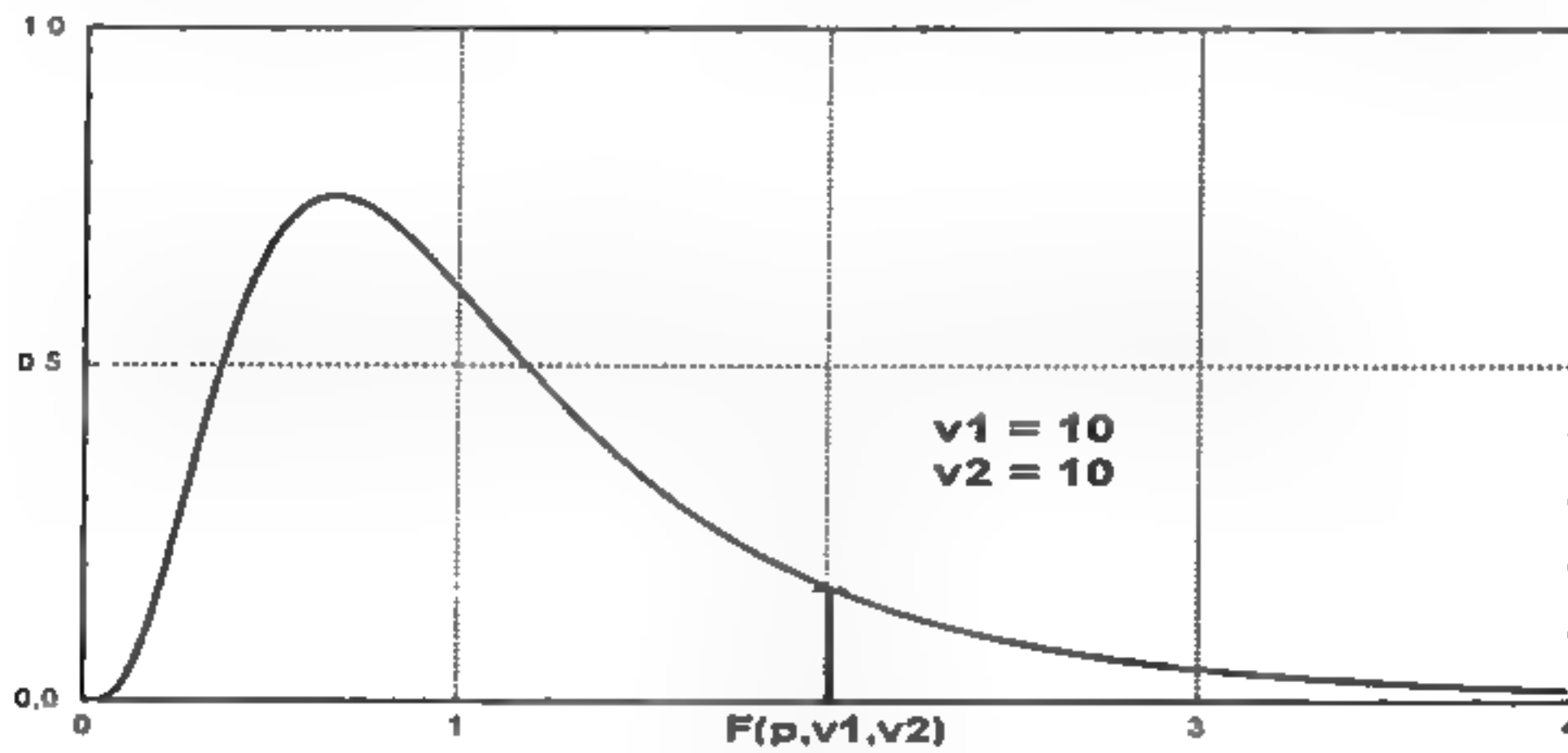
➤ مثال 2:  $(0.025, 10) = 3.247$

### 6.5.1- توزيع F (فيشر، سنيديكور، Fischer, Snedecor) :

أحد التوزيعات الإحصائية المتصلة المهمة والمستخدم في اختبار الفرضيات و في تحليل التباين. يعرف متغيره العشوائي بالدالة الإحصائية التالية:

$$f(F) = \frac{C \cdot F^{(v_1-2)/2}}{(v_2 + v_1 \cdot F)^{(v_1+v_2)/2}} \quad (1.38) \quad \text{و } F > 0$$

يرمز له بالرمز  $F(v_1, v_2)$ ، حيث  $v_1$  و  $v_2$  هما درجات الحرية و  $C$  هو ثابت يعتمد على  $v_1$  و  $v_2$  ليجعل المساحة تحت المنحنى يساوي 1.



إن لهذا التوزيع عدنان من درجة الحرية، وحيث أن  $v_2$  لا تظهر إلا في المقام، فإننا نسمي  $v_2$  درجات حرية المقام و نسمي  $v_1$  درجات حرية البسط. يقترب توزيع F من التوزيع الطبيعي بزيادة قيمتي  $v_1$  و  $v_2$ .



➤ مثال:

$$F(0.95, 9, 7) = 3.68$$

➤ ملاحظة: هناك بعض المساحات الصغيرة لا توجد في جداول توزيع F، ولإيجاد هذه المساحات فإنه يمكن استعمال القاعدة التالية:

$$F(p, v_1, v_2) = \frac{1}{F(1-p, v_2, v_1)} \quad (1.39)$$

➤ مثال:

$$F(0.05, 10, 7) = \frac{1}{F(0.95, 7, 10)} = \frac{1}{3.14} = 0.318$$

$$F(0.01, 11, 15) = \frac{1}{F(0.99, 15, 11)} = \frac{1}{4.25} = 0.23$$

• علاقة بين  $\chi^2$  و  $t$  و  $F$ :

$$F_{(1-p, 1, v)} = t^2_{(1-(p/2), v)} \quad (1.40) \quad (أ)$$

➤ مثال:

إذا كانت لدينا:  $p = 0.05$  و  $v = 3$  فإن:

$$F_{(1-0.05, 1, 3)} = 10.1$$

$$t_{(1-(p/2), v)} = t_{(1-(0.05/2), 3)} = 3.18$$

$$t^2 = 10.11$$

$$F_{(p, v, \infty)} = \frac{\chi^2_{(p, v)}}{v} \quad (1.41) \quad (ب)$$

➤ مثال:

$$F_{(0.95, 3, \infty)} = 2.60$$

$$\frac{\chi^2_{(0.95, 3)}}{3} = \frac{7.81}{3} = 2.60$$

## 6.1- التوزيعات الإحصائية الثنائية:

نفرض أن  $X$  و  $Y$  متغيران عشوائيان معرفان على نفس فضاء العينة  $S$  حيث:

$$S_X = \{Y_1, Y_2, \dots, Y_n\} \quad \text{و} \quad S_Y = \{X_1, X_2, \dots, X_n\}$$

إذا كان  $(X, Y)$  متغيرا عشوائيا ثنائيا فإن  $h(X, Y)$  يكون إقترانا احتماليا (توزيعا مشتركا) إذا تحقق ما يلي:

$$\begin{cases} h(x, y) \geq 0 \\ \sum_x \sum_y h(x, y) = 1 \end{cases} \quad (1.42)$$

$Y \backslash X$	$Y_1$	$Y_2$	.....	$Y_m$	المجموع
$X_1$	$h(x_1, y_1)$	$h(x_1, y_2)$	.....	$h(x_1, y_m)$	$f(x_1)$
$X_2$	$h(x_2, y_1)$	$h(x_2, y_2)$	.....	$h(x_2, y_m)$	$f(x_2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$X_n$	$h(x_n, y_1)$	$h(x_n, y_2)$	.....	$h(x_n, y_m)$	$f(x_n)$
المجموع	$g(y_1)$	$g(y_2)$	.....	$g(y_m)$	

نسمي:

$$f_x(x) = \sum_y h(x, y) \quad (1.43) \quad \text{بالكثافة الهامشية للمتغير } X$$

$$g_y(y) = \sum_x h(x, y) \quad (1.44) \quad \text{أما فهو إقتران الكثافة الهامشية لـ } Y.$$

أما الكثافة الشرطية للمتغير  $X$  إذا علم  $Y$  فهي:

$$h(x/y) = \frac{h(x, y)}{f_y(y)} \quad (1.45)$$

والعكس، الكثافة الشرطية للمتغير  $Y$  إذا علم  $X$  فهي:

$$h(y/x) = \frac{h(x, y)}{f_x(x)} \quad (1.46)$$

➤ مثال 1:

ليكن  $X$  المتغير العشوائي المعروف كما يلي:

$$\begin{cases} X=1 & \text{الشخص يدخن} \\ X=0 & \text{الشخص لا يدخن} \end{cases}$$

وليكن  $Y$  المتغير العشوائي المعروف كما يلي:

$$\begin{cases} X=1 & \text{الشخص مصاب بالسرطان} \\ Y=0 & \text{الشخص غير مصاب} \end{cases}$$

نلخص الإقتران الإحتمالي  $h(X,Y)$  كما يلي:

$Y$	0	1
$X$		
0	0.40	0.02
1	0.03	0.55

إن احتمال أن يكون الشخص وغي مصاب هو 0.40 أي:

$$P(X=0, Y=0) = h(0,0) = 0.40$$

$$P(X=0, Y=1) = h(0,1) = 0.02$$

.....

➤ مثال 2:

من خلال الإقتران الإحتمالي التالي يمكن إيجاد:

$Y$	0	1	$X f_x(x) = \sum_y h(x,y)$
$X$			
0	0.1	0.0	0.1
1	0.3	0.4	0.7
2	0.1	0.1	0.2
الكثافة الهامشية $Y$ $g_Y(y) = \sum_x h(x,y)$	0.5	0.5	1

أما الإحتمالات الشرطية  $h(x/y) = \frac{h(x,y)}{f_y(y)}$  فتكون على الشكل التالي:

X \ Y	0	1
	0	1
0	0.2	0
1	0.6	0.8
2	0.2	0.2

$$h(x/y) = \frac{h(x,y)}{f_y(y)} = f(x/y) = \frac{h(0,1)}{0.5} = \frac{0.1}{0.5} = 0.2$$

$$h(x/y) = \frac{h(x,y)}{f_y(y)} = f(x/y) = \frac{h(1,1)}{0.5} = \frac{0.3}{0.5} = 0.6$$

$$h(x/y) = \frac{h(x,y)}{f_y(y)} = f(x/y) = \frac{h(1,2)}{0.5} = \frac{0.1}{0.5} = 0.2$$

.....

➤ ملاحظة:

يكون X و Y مستقلين إذا كان:

$$(1.44) h(X,Y) = f_x(X).g_y(Y)$$

ففي المثال السابق نجد أن:

$$h(X,Y) \neq f_x(X).g_y(Y)$$

أي أن المتغيرين مرتبطان.  $h(0,1) \neq 0.1 \times 0.5$

$$0.1 \neq 0.05$$

### 1.6.1- الأمل الرياضي:

(أ) - حالة المتغيرات العشوائية المنفصلة:

إذا كان  $X$  و  $Y$  متغيران عشوائيان منفصلين و معرفين بالإقتران  $h(X, Y)$  فإن:

$$E(X) = \mu_x = \sum_x \sum_y X \cdot h(X, Y) \quad (1.47)$$

$$E(Y) = \mu_y = \sum_y \sum_x Y \cdot h(X, Y) \quad (1.46)$$

(ب) - حالة المتغيرات العشوائية المتصلة:

إذا كان  $X$  و  $Y$  متغيران عشوائيان متصلين و معرفين بالإقتران  $f(X, Y)$  فإن:

$$E(X) = \mu_x = \iint_R X \cdot h(X, Y) dX dY \quad (1.48)$$

$$E(Y) = \mu_y = \iint_R Y \cdot h(X, Y) dX dY \quad (1.49)$$

### 2.6.1- التباين:

(أ) - حالة المتغيرات العشوائية المنفصلة:

$$\sigma^2(X) = \sum_x \sum_y (X - E(X))^2 \cdot h(X, Y) \quad (1.50)$$

$$\sigma^2(Y) = \sum_x \sum_y (Y - E(Y))^2 \cdot h(X, Y) \quad (1.51)$$

(ب) - حالة المتغيرات العشوائية المتصلة:

$$\sigma^2(X) = \iint_R (X - E(X))^2 \cdot h(X, Y) dX dY \quad (1.52)$$

$$\sigma^2(Y) = \iint_R (Y - E(Y))^2 \cdot h(X, Y) dX dY \quad (1.53)$$

### 3.6.1- التباين المشترك (التغاير):

له بالرمز  $Cov(X, Y)$ ، و يعطى بالعلاقة التالية:

ويرمز (أ) - حالة المتغيرات العشوائية المنفصلة:

$$Cov(X, Y) = \sigma_{xy} = \sum_i \sum_j (X_i - \mu_x)(Y_j - \mu_y) \cdot h(X, Y) \quad (1.54)$$

ويمكن تبسيطه ليأخذ الشكل التالي:

$$\text{Cov}(X, Y) = \sigma_{xy} = \sum_i \sum_j X_i Y_j h(X, Y) - \mu_x \mu_y \quad (1.55)$$

$$= E(X, Y) - E(X) \cdot E(Y)$$

(ب) - حالة المتغيرات العشوائية المتصلة:

$$\text{Cov}(X, Y) = \iint_R (X_i - \mu_x)(Y_i - \mu_y) \cdot h(X, Y) dX dY \quad (1.56)$$

#### 4.6.1- معامل الارتباط:

يقيس معامل الارتباط شدة العلاقة بين المتغير  $X$  و المتغير  $Y$ ، و يعطى بالعلاقة التالية:

$$\rho(x, y) = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y} \quad (1.57)$$

حيث:  $0 \leq \rho \leq 1$

إذا كان  $\text{Cov}(x, y) = 0$  فإن  $\rho = 0$ ، أي أن  $X$  و  $Y$  مستقلان.

➤ تمرين:

نفترض التوزيع المشترك الاحتمالي التالي:

	Y	-3	2	4	المجموع
X					
1		0.1	0.2	0.2	0.5
2		0.3	0.1	0.1	0.5
المجموع		0.4	0.3	0.3	1

(أ): أحسب التباين المشترك لهذا التوزيع،

(ب): أحسب معامل الارتباط، هل  $X$  و  $Y$  مستقلان؟.

➤ الحل:

لدينا:

$$\text{Cov}(X, Y) = \sigma_{xy} = \sum_i \sum_j X_i Y_j h(X, Y) - \mu_x \mu_y$$

$$\mu_x = \sum X f(X) = 1 \times 0.5 + 3 \times 0.5 = 2$$

$$\mu_y = \sum Y g(Y) = -3 \times 0.4 + 2 \times 0.3 + 4 \times 0.3 = 0.6$$

$$E(X, Y) = \sum_i \sum_j X_i Y_j h(X, Y) = 1 \times (-3) \times 0.1 + 1 \times 2 \times 0.2 + \dots + 3 \times 4 \times 0.1 = 0$$

ومنه:

$$\text{Cov}(X, Y) = 0 - 2 \times 0.6 = -1.2$$

من جهة أخرى لدينا:

$$\sigma^2_x = \sum X^2 \cdot f(X) - \mu_x^2 = (1 \times 0.5 + 9 \times 0.5) - 4 = 1$$

$$\sigma^2_y = \sum X^2 \cdot g(X) - \mu_y^2 = (9 \times 0.4 + 4 \times 0.3 + 16 \times 0.3) - 0.36 = 9.24$$

لنصل إلى:

$$\rho(x, y) = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y} = \frac{-1.2}{\sqrt{1} \times \sqrt{9.24}} = -0.4$$

من المؤكد أن  $X$  و  $Y$  غير مستقلين لأن معامل الارتباط غير معدوم، ولكن يمكن التأكد من العلاقة التالية:

$$h(X, Y) = f_x(X) \cdot g_y(Y)$$

$$h(1, -3) \neq 0.4 \times 0.5$$

$$0.1 \neq 0.2$$

## تمارين مختارة

### التمرين الأول:

صنعت قطعة نقود بحيث يكون احتمال ظهور الصورة هو ضعف احتمال ظهور الكتابة في الرمية الواحدة. أوجد احتمال ظهور الصورة، أوجد احتمال ظهور الكتابة.

### التمرين الثاني:

إختيرت 03 مصابيح كهربائية بطريقة عشوائية من بين 15 مصباح كهربائي، 05 منها فاسدة. أوجد الإحتمال  $P$  بحيث يكون:  
- جميعها سليمة، - واحدة فقط فاسدة، - واحدة على الأقل فاسدة.

### التمرين الثالث:

إذا كان احتمال إن يصيب 03 رجال هدفا هو على التوالي:  $1/4$ ،  $1/6$ ،  $1/3$ . فإذا كان كل منهم يصوب مرة واحدة على الهدف، أوجد احتمال: - أن يصيب الهدف مرة واحدة. - إذا أصاب الهدف رجل واحد فقط، ما هو احتمال أن يكون الرامي الثاني؟.

### التمرين الرابع:

في إحدى الجامعات، وجد أن 4% من الطلبة و 1% من الطالبات أطولهم أكثر من 1.80م وأن 60% من مجموع طلبة الجامعة إناث. أختير بطريقة عشوائية أحد الطلبة ووجد أن طوله أكثر من 1.80م. ما هو احتمال أن يكون هذا الإختيار طالبة.

### التمرين الخامس:

تنتج 3 ماكينات على التوالي 60%، 30%، 10% من الإنتاج الكلي لمصنع ما؛ فإذا كانت نسبة الإنتاج الفاسد لهذا المصنع هو على التوالي: 2%،



3 %، 4 % . أختيرت وحدة بطريقة عشوائية ووجدت أنها فاسدة. أوجد احتمال أن تكون من الماكينة الثالثة.

التمرين السادس:

ترمى زهرة نرد مرة واحدة. نعرف  $X$  المتغير العشوائي الذي يمثل ضعف العدد الذي يظهر.

- أنشئ جدول قانون التوزيع الإحتمالي المقابل ثم مثل بيانيا هذا التوزيع،
- أحسب أمله الرياضي ثم إنحرافه المعياري.

التمرين السابع:

ليكن  $X$  متغير عشوائي متصل معرفة بالدالة الإحتمالية التالية:

$$f(X) = \begin{cases} 1/2 & 1 \leq X \leq 3 \\ 0 & \text{دون ذلك} \end{cases}$$

- أحسب  $P(0 \leq X \leq 3)$  ، - أحسب الأمل الرياضي ثم الإنحراف المعياري،
- مثل بيانيا هذا التوزيع.

التمرين الثامن:

إذا كان 10 % من الطلبة الأوائل ينتقلون من المدى القصير إلى المدى الطويل، فإذا كانت علامات الطلبة موزعة طبيعيا بمتوسط 72 وإنحراف معياري 9؛ فما هي علامة تأخذ بعين الاعتبار لإنجاز عملية الانتقال ؟.

التمرين التاسع:

تم حقن مريض بمضاد حيوي معين، فإذا كان احتمال أن يسبب المضاد الحيوي حساسية هو 0.01، فأوجد الاحتمال أنه من بين 2000 مريض تم حقنهم بالمضاد الحيوي أن يكون:

- مريض واحد يصاب بالحساسية، - 3 مرضى سيصابون،
- أكثر من 3 مرضى سيصابون بالحساسية.

### التمرين العاشر:

متوسط طول 500 ورقة نبات ما هو 151 مم بإنحراف معياري 15 مم، فإذا كانت الأطوال موزعة طبيعيا، فما هو عدد الأوراق التي أطواها:  
- بين 120 و 155 مم، - أكبر من 185 مم.

### التمرين الحادي عشر:

إذا كانت علامات الطلبة في مقياس الإحصاء الرياضي موزعة طبيعيا وكانت:

$$\sigma = 2 \sum X^2 = 14800, n = 100,$$

- إحصاء احتمال أن تكون علامة الطالب تزيد عن 17،
- إحصاء احتمال أن تكون علامة الطالب تقل عن 17.

### التمرين الثاني عشر:

ألقيت قطعة نقود 12 مرة، ما هو احتمال أن نحصل على الصورة عدد من المرات يتراوح بين 4 و 7 و ذلك باستخدام:  
- توزيع ذو الحدين،  
- التقريب الطبيعي لذو الحدين.

### التمرين الثالث عشر:

ألقيت قطعة نقود 3 مرات. نفرض  $X$  الذي يدل على 0 أو 1 تبعا لظهور الصورة أو الكتابة في الرمية الأولى؛ و نفرض أن  $Y$  تدل على عدد مرات ظهور الصورة.  
- أوجد توزيع المتغير  $X$  و المتغير  $Y$ ،  
- التوزيع المشترك لـ  $X$  و  $Y$ ،  
- أحسب التباين المشترك لهذا التوزيع ثم معلل الارتباط،  
- هل  $X$  و  $Y$  مستقلين؟.

### التمرين الرابع عشر:

ليكن  $X$  المتغير العشوائي و له التوزيع التالي:

$X:$        -2     -1     1     2

$F(X):$    1/4    1/4   1/4    1/4

وأن  $Y = X^2$

- أوجد التوزيع  $g$  الموافق للمتغير  $Y$ ،
- أوجد التوزيع المشترك  $h(X, Y)$ ،
- معامل الارتباط لهذا التوزيع.

### التمرين الخامس عشر:

ليكن  $X$  متغير عشوائي معرف من خلال رمي قطعة نقود متزنة حيث  $X = 1$  عند ظهور الصورة،  $X = 2$  عندما تظهر الكتابة. ترمى قطعة النقود 3 مرات متتالية، وليكن  $X_1$  للرمية الأولى،  $X_2$  للرمية الثانية،  $X_3$  للرمية الثالثة. نعرف المتغير  $Y$  بالقانون:  $Y = X_1^2 - X_2 - X_3$ .

- أوجد القيم الممكنة  $y_1, y_2, \dots$  الموافقة لـ  $Y$  وماهي الاحتمالات المقابلة  $P_1, P_2, \dots$
- مثل بيانيا دالة التوزيع الاحتمالية  $f(y)$ ، حيث  $f(Y < y)$
- أحسب  $E(Y)$  ثم  $\sigma$ .



## الفصل الثاني

### ختبار الفرضيات

مقدمة:

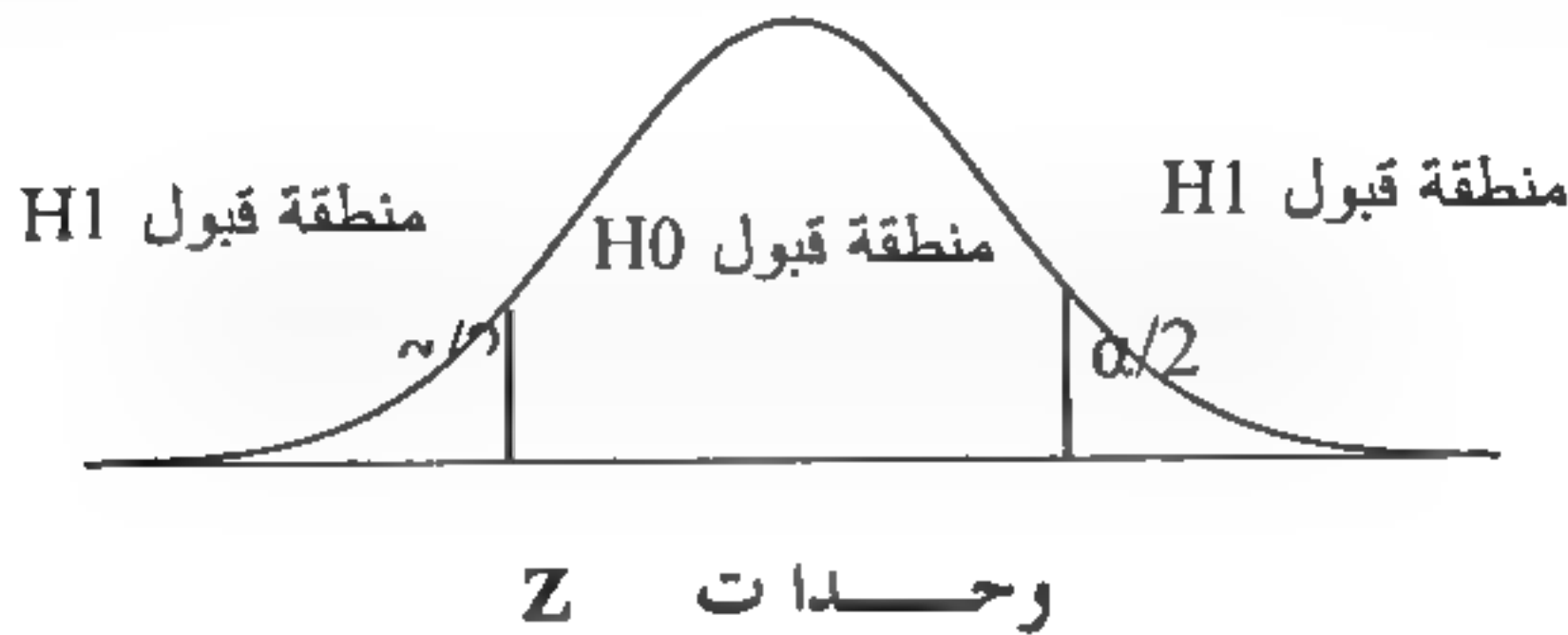
الفرضية هي جملة حول مجتمع إحصائي أو أكثر معالمه  $(\sigma, \mu)$  مأخوذة منه عينة عشوائية، وعلى أساس القرار الإحصائي، إما أن نقبل وإما أن نرفض هذه الفرضية بدرجة ثقة معلومة (إحتمال القبول أو الرفض). يتم فحص الفرضية وفق الخطوات التالية:

- دراسة طبيعة البيانات الإحصائية، هل هي نوعية، متصلة..... بغرض معرفة طبيعة التوزيع الذي تخضع له البيانات،

- وضع الفرضيات، ففي الغالب يوجد نوعين من الفرضيات: الفرضية العدمية  $H_0$  التي لا تعبر عن وجود تغيرات والفرضية البديلة  $H_1$  التي تمثل عكس الفرضية العدمية،

- دالة الاختبار، والتي تساعد في إتخاذ القرار حول الفرضية، فالدالة هي متغير عشوائي، لأن قيمتها تتغير بتغير العينة الإحصائية لذلك يجب معرفة طبيعة التوزيع الإحتمالي الذي تخضع له البيانات،

- القرار الإحصائي، ففي الواقع ينقسم مجال دالة الاختبار إلى منطقتين، منطقة قبول  $H_0$  وهو ما يعني رفض  $H_1$  ومنطقة رفض  $H_0$  وهو ما يعني قبول  $H_1$



إن إتخاذ أي قرار إحصائي ينطوي على خطأ بنسبة ما، فمن المحتمل أن نرفض فرضية لإي حين أنها صحيحة أو العكس، ولذلك يمكن أن نستنتج نوعين من الأخطاء. الخطأ من النوع الأول و الناتج عن رفض  $H_0$  عندما تكون هذه الأخيرة صحيحة، و الخطأ من النوع الثاني و الناتج عن قبول  $H_0$  في حين أنها خاطئة. يكون احتمال الخطأ الأول هو  $\alpha$ ، أي أن مستوى الثقة يساوي  $1-\alpha$ ، أما خطأ النوع الثاني فيرمز له بالرمز  $\beta$  (إحتمال فرض خاطئ)، فإذا نقصت قيمة  $\alpha$  فسوف نضطر إلى قبول احتمال أكبر لـ  $\beta$ ، ولتفادي ذلك لإغنه يجب زيادة حجم العينة. ويمكن تلخيص ذلك في الجدول التالي:

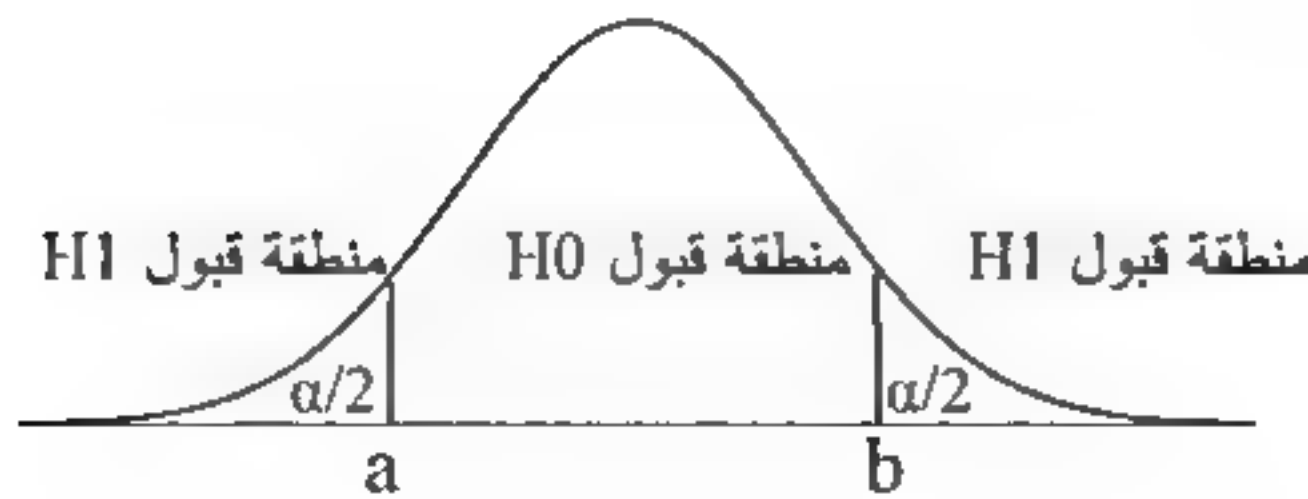
الحقيقة \ القرار	$H_0$ صحيحة	$H_0$ خاطئة
قبول $H_0$	قرار صحيح	خطأ من النوع الثاني
رفض $H_0$	خطأ من النوع الأول	قرار صحيح

- الخطأ من النوع الأول  $\alpha$  هو احتمال رفض  $H_0$  في حين أنها صحيحة،
- الخطأ من النوع الثاني  $\beta$  احتمال قبول  $H_0$  في حين أنها خاطئة.

## 1.2- أنواع الفرضيات:

- الفرضية البديلة ذات الذيلين:
- وتكون على الشكل التالي:

$$\begin{cases} H_1: \mu \neq \mu_0 \\ H_1: \sigma^2 \neq \sigma_0^2 \end{cases}$$



فإذا كانت  $G$  هي دالة الاختبار فإن:

$$P(G < a) = p(G > b) = \alpha/2 \quad (2.1)$$

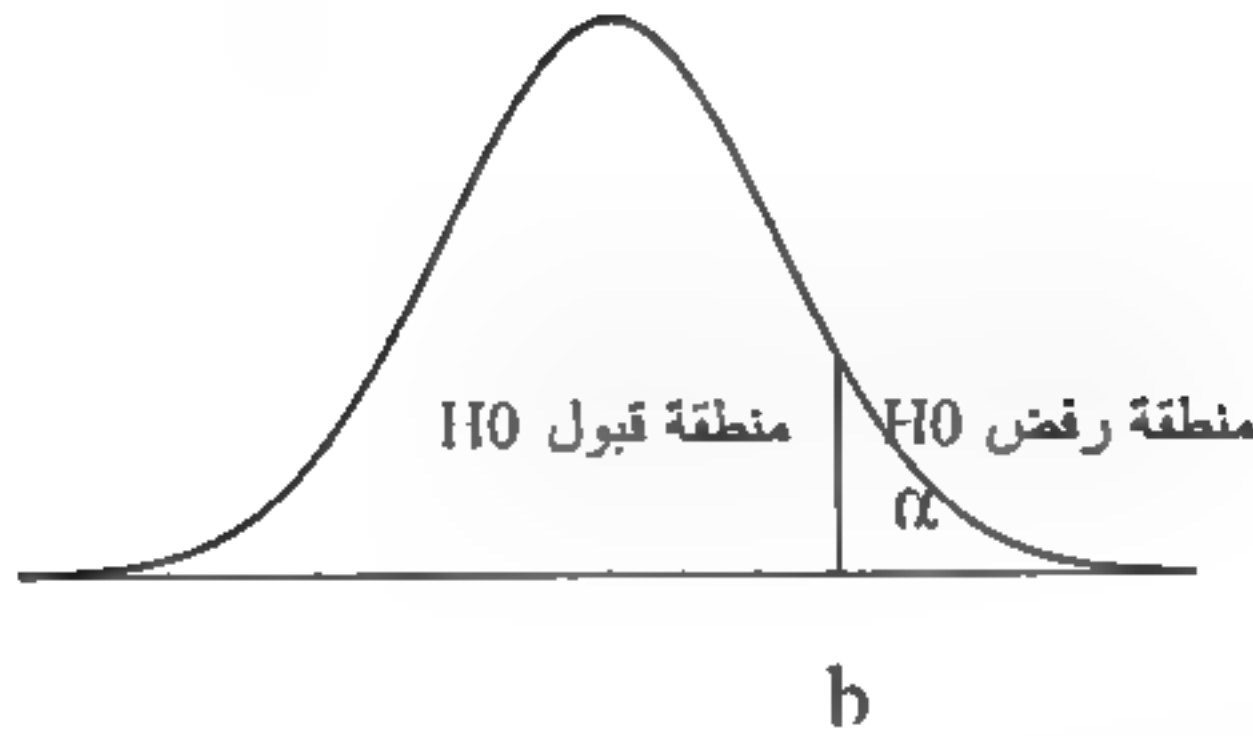
• الفرضية البديلة ذات الذيل الأعلى:

وتكون على الشكل التالي:

$$\begin{cases} H_1: \mu > \mu_0 \\ H_1: \sigma^2 > \sigma_0^2 \end{cases}$$

في هذه الحالة تكون  $\alpha$  في الطرف الأعلى من توزيع الدالة  $G$  ويكون:

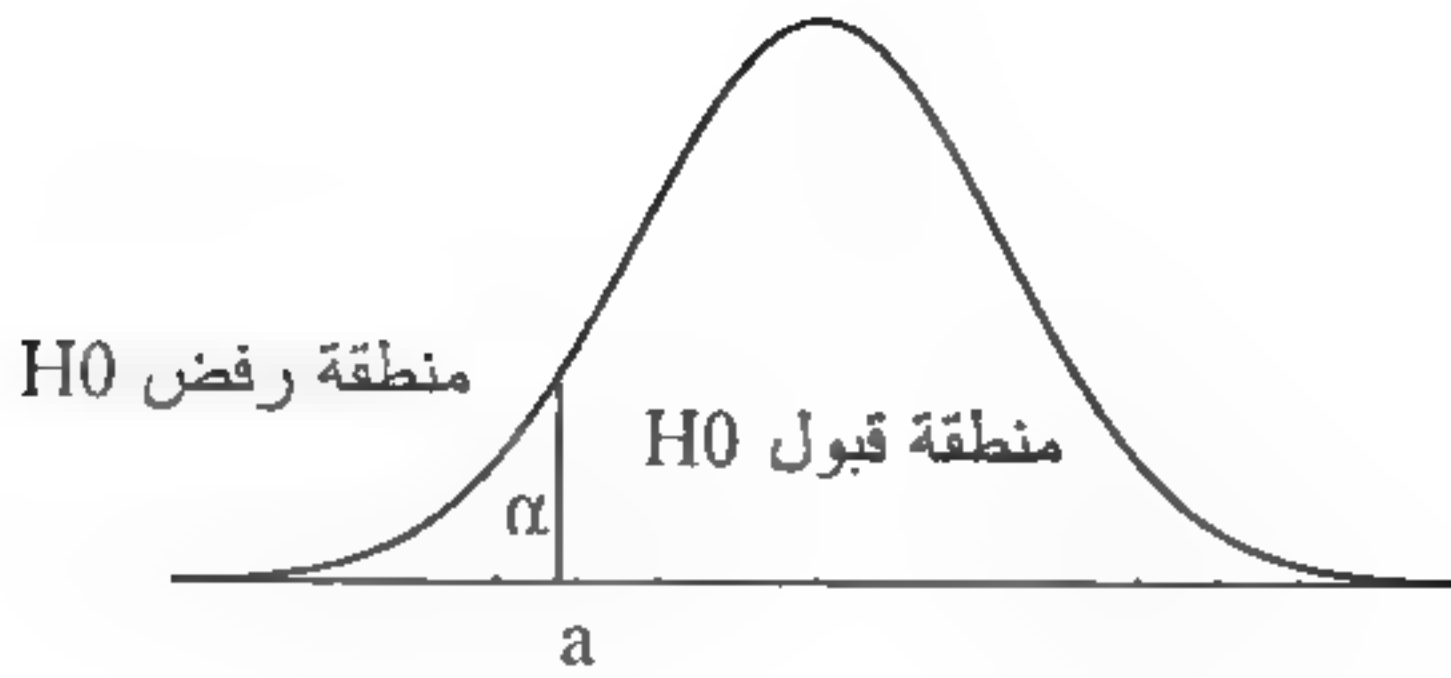
$$p(G > b) = \alpha \quad (2.2)$$



• الفرضية البديلة ذات الذيل الأدنى

$$\begin{cases} H_1: \mu < \mu_0 \\ H_1: \sigma^2 < \sigma_0^2 \end{cases}$$

$$p(G < a) = \alpha \quad (2.3)$$



## 2.2- اختبار الأوساط الحسابية:

لتكن  $\mu$  هو المتوسط الحسابي لمجتمع ما، ولتكن  $\mu_0$  قيمة إفتراضية لـ  $\mu$ . نريد أن نختبر عند مستوى دلالة  $\alpha$  إذا ما كان:

$$\begin{aligned} H_0: & \mu = \mu_0 \\ H_1: & \mu \neq \mu_0 \end{aligned}$$

$$Z = \frac{\bar{X} - \mu_0}{\sigma_{\bar{x}}} \quad (2.4) \quad \text{وتكون دالة الاختبار:}$$

➤ مثال:

تنتج شركة كهربائية مصابيح كهربائية، وترغب هذه الشركة في معرفة إذا ما كان يمكنها الإدعاء بأن مصابيحها الكهربائية تستمر 1000 ساعة دون احتراق. أخذنا عينة عشوائية من هذه المصابيح حجمها 100 مصباح ووجدنا أن المتوسط الحسابي للمصابيح دون أن تحترق هو 980 ساعة بإنحراف معياري يساوي 80 ساعة. إختبر إذا ما كان إدعاء الشركة صحيحا عند  $\alpha = 5\%$ .

$$\begin{aligned} H_0: & \mu = 1000 \\ H_1: & \mu \neq 1000 \end{aligned}$$



وتكون دالة الاختبار:

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{X} - \mu_0}{\frac{80}{\sqrt{100}}} = \frac{980 - 1000}{\frac{80}{\sqrt{100}}} = -2.50$$



وحيث أن قيمة  $Z$  المحسوبة وقعت في منطقة رفض  $H_0$  فإن عمر المصابيح الكهربائية لا يساوي 1000 ساعة.

➤ مثال 2:

ترغب شركة أن تعرف بدرجة ثقة 95% إذا ما كان يمكنها الإدعاء بأن صناديق الصابون المسحوق الذي تنتجه تحتوي على أكثر من 500 غ من المادة الفعالة. إذا كانت الأوزان تخضع للتوزيع الطبيعي، أخذنا عينة عشوائية حجمها 25 وحدة ووجدنا أن المتوسط يساوي 520 غ، بانحراف معياري يساوي 50 غ. إختبر إذا ما كان المتوسط الحسابي للمجتمع أكبر من 500 غ.

➤ الحل:

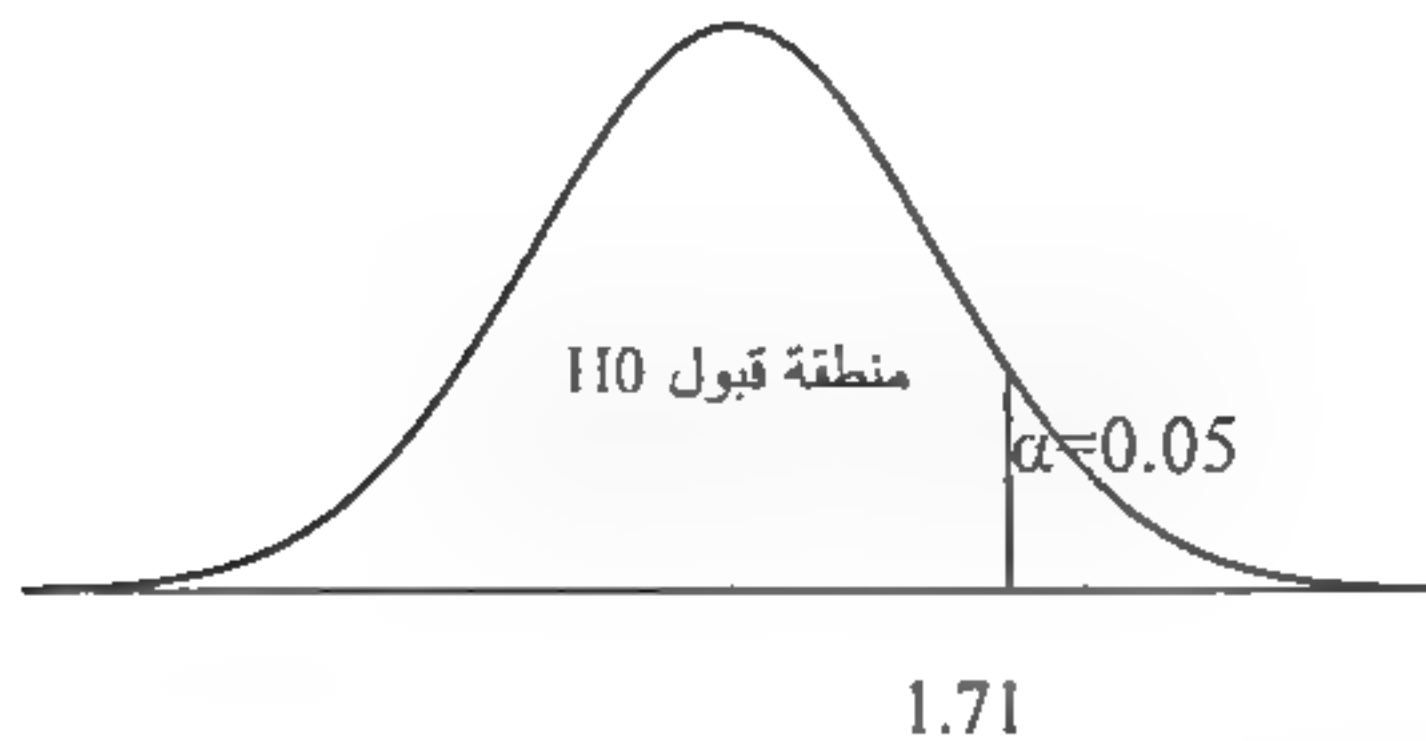
$$\begin{aligned} H_0: \mu &= 500 \\ H_1: \mu &> 500 \end{aligned}$$

لدينا الفرضيتان التاليتان:

وحيث أن  $n < 30$  و  $\sigma$  مجهولة فإن دالة الاختبار تخضع لتوزيع  $t$  بدرجات حرية 24، ويكون  $s$  تقديرا لـ  $\sigma$ :

$$t = \frac{\bar{X} - \mu}{s_{\bar{x}}} = \frac{520 - 500}{75/\sqrt{25}} = 1.33$$

وحيث أن  $t$  التطبيقية أقل من  $t$  النظرية (وقعت في منطقة قبول  $H_0$ )، فإن مسحوق الصابون تحتوي على 500 غ من المادة الفعالة.



## 1.2.2 - اختبار الفرضيات للفرق بين وسطين:

ليكن  $\bar{X}_1$  و  $\bar{X}_2$  متوسطين حسابيين لعينتين حجمها على التوالي  $n_1$  و  $n_2$  حيث:  $n_1$  و  $n_2$  أكبر من 30، مأخوذتين من مجتمعين وسطهما على التوالي  $\mu_1$  و  $\mu_2$ ، بانحرافات معيارية  $\sigma_1$  و  $\sigma_2$ . لإختبار إذا ما كان:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

فإن دالة الاختبار:

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}} \quad (2.5)$$

حيث:

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad (2.6)$$

عند مستوى دلالة  $\alpha$ .

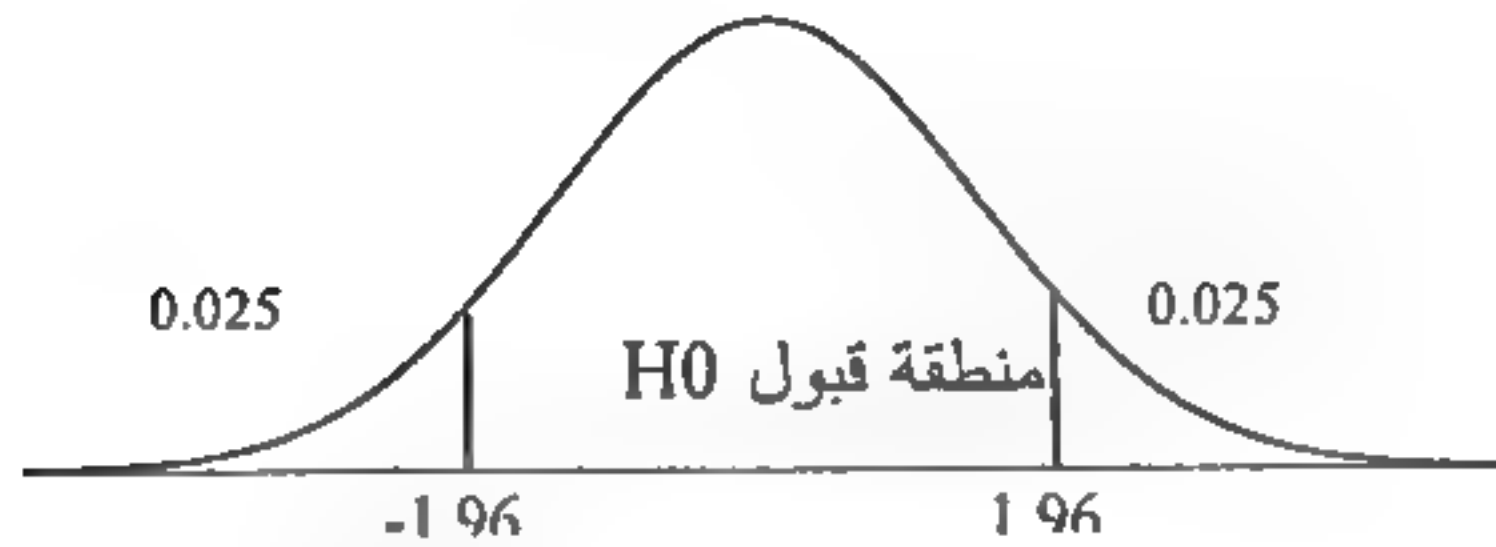
### ➤ مثال:

يرغب مدير مؤسسة أن يحدد عند مستوى دلالة 5% إذا ما كان الأجر بالساعة للعمال متساويا في مدينتين. أخذنا عينتين عشوائيتين من المدينتين حجمهما على التوالي 40 عاملا و 54 عاملا، ووجدنا المتوسط الحسابي للعينة الأولى يساوي 6 دج للساعة وللعينة الثانية هو 5.40 دج للساعة، بانحرافات معيارية هي على التوالي 2 دج و 1.80 دج. إختبر إذا ما كانت هناك فروق معنوية بين أجور العمال في المدينتين.

### ➤ الحل: لدينا الفرضيات التالية:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$



$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{6 - 5.40}{\sqrt{\frac{2^2}{40} + \frac{1.80^2}{54}}} = 1.5$$

نلاحظ أن قيمة Z النظرية وقعت داخل منطقة قبول  $H_0$  وعليه لا توجد فروق معنوية بين المدينتين.

• إذا كانت  $n_1$  و  $n_2$  أقل من 30:

تخضع دالة الاختبار في هذه الحالة إلى توزيع  $t$  بدرجات حرية  $n_1 + n_2 - 2$  و يسمى باختبار العينات الصغيرة، بحيث نفترض أن:

$\sigma_1 = \sigma_2 = \sigma$  ويكون  $s_1$  و  $s_2$  تقديرين لـ  $\sigma_1$  و  $\sigma_2$  لأنهما غير معلومتين و  $s$  تقديرا لـ  $\sigma$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (2.7)$$

حيث:

$$s = \sigma = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \quad (2.8)$$

➤ مثال:

ترغب إدارة الجامعة في معرفة إذا ما كانت هناك فروقا معنوية في نتائج الطلبة المسجلين بجامعة مع نتائج طلبة جامعة أخرى، في نفس التخصص. أخذت عيّنتين عشوائيتين حجمهما، 21 طالبا لكل جامعة، وكانت الأوساط الحسابية على التوالي: 78 بانحراف معياري قدره 8 و 74 بانحراف معياري قدره 7. هل توجد فروق معنوية في علامات الطلبة؟.

➤ الحل:

لدينا الفرضيات التالية:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

نقوم بحساب قيمة  $t$  التطبيقية، حيث:

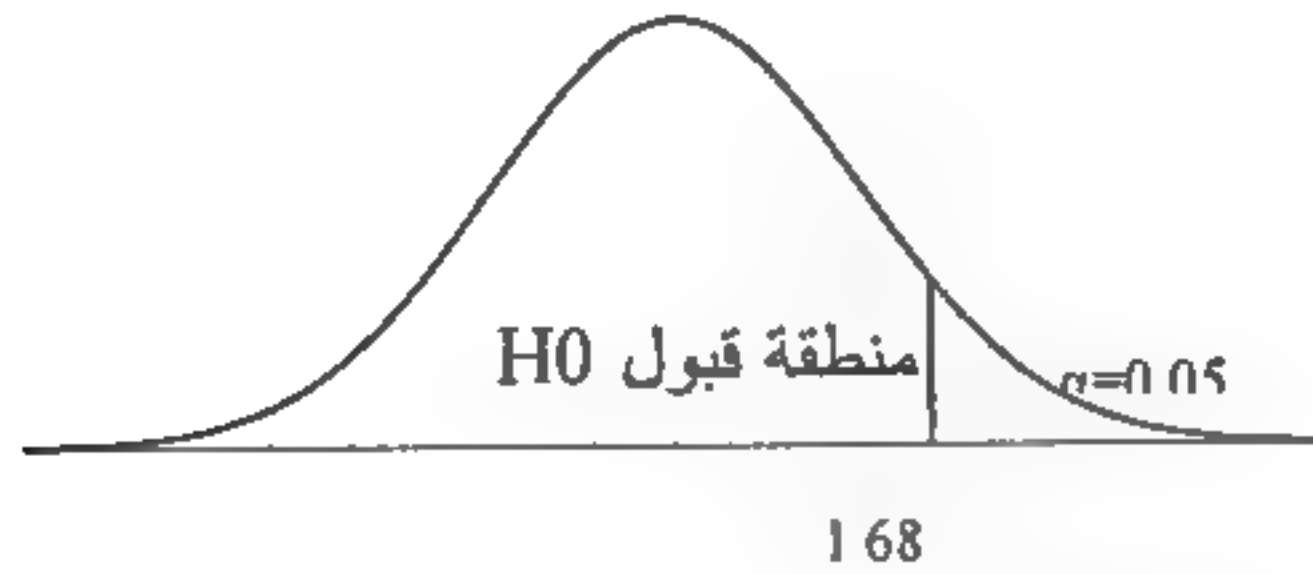
$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{78-74}{7.51 \sqrt{\frac{1}{21} + \frac{1}{21}}} = 1.72$$

$$s = \sigma = \sqrt{\frac{(21-1)64 + (21-1)49}{40}} = 7.51 \text{ مع:}$$

إن قيمة  $t$  النظرية يمكن حسابها من جدول  $t$  بدرجات حرية 40 وعند المساحة 0.95 (أو  $\alpha=0.05$ )، فنجد:

$$t_{(0.95,40)}=1.68$$

نلاحظ أن قيمة  $t$  التطبيقية وقعت خارج منطقة قبول  $H_0$  وبالتالي هناك



فروق معنوية بين الجامعتين.

### 3.2- إختبار التباين:

لنفترض أنه لدينا مجتمع موزع طبيعياً، بانحراف معياري  $\sigma$ . نريد أن نختبر إذا ما كان  $\sigma = \sigma_0$  مقابل  $\sigma > \sigma_0$  عند مستوى دلالة  $\alpha$ . إن هذا الإختبار يخضع لتوزيع كاي تربيع بدرجات حرية  $n-1$ ، حيث أن:

$$\chi^2 = \frac{ns^2}{\sigma^2} \quad (2.9)$$

حيث  $n$  هي حجم العينة و  $s^2$  هو تباين هذه العينة.

➤ مثال:

من الخبرة الماضية وجدنا أن الانحراف المعياري لمجتمع ما هو 0.25. أخذنا عينة عشوائية حجمها 20 من هذا المجتمع ووجدنا أن إنحرافها المعياري يساوي 0.32. هل لهذا الارتفاع في الارتفاع دلالة عند مستوى معنوية 0.05.

➤ الحل:

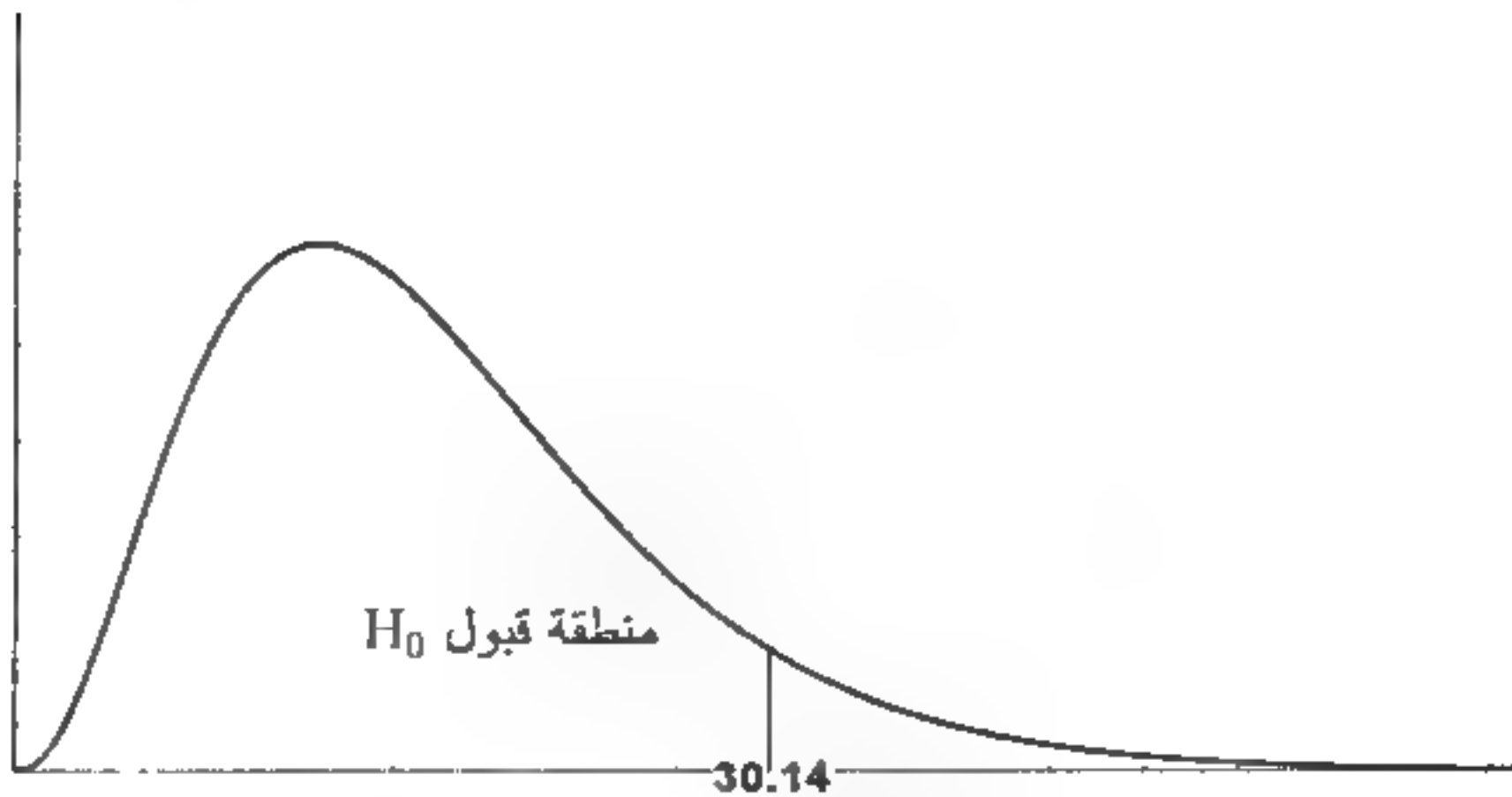
لدينا الفرضيات التالية:

$$H_0: \sigma = 0.25$$

$$H_1: \sigma > 0.25$$

$$= \frac{20 \times 0.32^2}{0.25^2} = 32.8 \chi^2 = \frac{ns^2}{\sigma^2}$$

وحيث أن  $\chi^2$  المحسوبة أكبر من  $\chi^2$  النظرية (30.14)، فإننا نرفض الفرضية العدمية وعليه فإنه يوجد زيادة في الانحراف المعياري.



#### 4.2- اختبار النسبة بين تباينين:

ليكن لدينا مجتمعان موزعين طبيعيا، أخذنا عينتين عشوائيتين منهما حجمهما على التوالي  $n$  و  $m$ ، بانحرافات معيارية، على التوالي:  $s_1^2$  و  $s_2^2$ . نريد إجراء الاختبار التالي:

$$H_0: \sigma_1 = \sigma_2$$

$$H_1: \sigma_1 > \sigma_2$$

إن هذا الاختبار يخضع لتوزيع  $F$ ، وفق العلاقة التالية:

$$F = \frac{\hat{S}_1^2}{\hat{S}_2^2} \quad (2.10)$$

$$\hat{S}_2^2 = \frac{n_2}{n_2 - 1} S_2^2 \quad \text{و} \quad \hat{S}_1^2 = \frac{n_1}{n_1 - 1} S_1^2 \quad \text{حيث:}$$

➤ مثال:

يعطي أستاذ دروسه لمجموعتين من الطلبة، المجموعة الأولى حجمها 16 طالبا و المجموعة الثانية حجمها 25 طالبا. بعد إجراء الإمتحان وجدنا أن الانحراف المعياري للمجموعة الأولى هو 9 وللمجموعة الثانية هو 12. هل يوجد فروق معنوية بين التباينين عند مستوى معنوية 0.05.

➤ الحل:

لدينا الاختبار التالي:

$$H_0: \sigma_1 = \sigma_2$$

$$H_1: \sigma_1 > \sigma_2$$

من جهة أخرى:

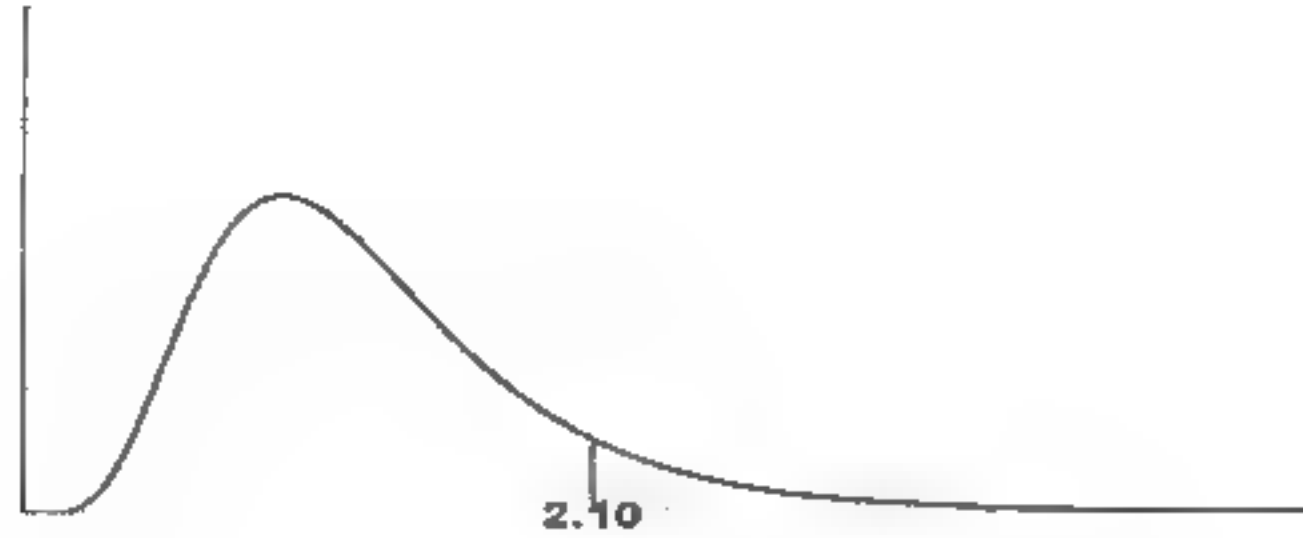
$$\hat{S}_1^2 = \frac{n_1}{n_1 - 1} S_1^2 = \frac{16}{15} \cdot 81 = 86.4$$

$$\hat{S}_2^2 = \frac{n_2}{n_2 - 1} S_2^2 = \frac{25}{24} 144 = 150$$

ومنه فإن:

$$F = \frac{\hat{S}_1^2}{\hat{S}_2^2} = \frac{86.4}{150} = 0.58$$

من جدول توزيع F يمكن حساب القيمة النظرية عند المساحة 0.95 ودرجات حرية  $v_1 = 15$  و  $v_2 = 24$  والتي تساوي 2.10. وحيث أن قيمة F النظرية وقعت داخل منطقة قبول  $H_0$  فإنه لا توجد فروق معنوية بين المجموعتين.



المجموعتين.

## 5.2- مقارنة بين نسبة نظرية و نسبة مشاهدة:

تعتمد المقارنة بين النسبة النظرية  $p_0$  لعينة مأخوذة من مجتمع والنسبة المشاهدة  $p$  لنفس العينة، على حساب قيمة  $\varepsilon$  التي تعطى بالعلاقة التالية:

$$\varepsilon = \frac{|p - p_0|}{\sqrt{\frac{p_0 q_0}{n}}} \quad (2.11)$$

عند مستوى دلالة 0.05،

فإذا كان:

- $\varepsilon < 1,96$ : فإنه لا توجد فروق معنوية بين النسبتين عند مستوى معنوية 5%
- $\varepsilon \geq 1,96$ : توجد فروق معنوية بين النسبتين عند مستوى معنوية 5%
- $\varepsilon < 2,576$ : فإنه لا توجد فروق معنوية بين النسبتين عند مستوى معنوية 1%
- $\varepsilon \geq 2,576$ : توجد فروق معنوية بين النسبتين عند مستوى معنوية 1%



➤ مثال:

في مجتمع سكاني، 20% مصابون بسرطان الرئة، أخذنا عينة عشوائية من هذا المجتمع حجمها 20 شخص، ووجدنا أن نسبة الإصابة بلغت 34%. هل توجد فروق معنوية بين النسبة المشاهدة والنسبة النظرية عند مستوى دلالة 0.05.

➤ الحل:

يمكننا حساب قيمة  $\varepsilon$  وفق العلاقة (2.11):

$$\varepsilon = \frac{|0,34 - 0,20|}{\sqrt{\frac{0,2 \times 0,8}{20}}} = 1,55 < 1,96$$

وهو ما يعني لا وجود لفروق معنوية بين النسبتين عند مستوى دلالة 0.05.

ملاحظة: يمكن إجراء المقارنة باستعمال  $\chi^2$ ، نتحصل على الجدول التالي:

	نسبة المصابين	نسبة الغير مصابين	المجموع
$C_i$ النسب النظرية	20 %	80 %	100 %
$O_i$ النسب المشاهدة	34 %	66 %	100 %

$$\chi^2 = \frac{\sum (O_i - C_i)^2}{C_i} = \frac{(34 - 20)^2}{20} + \frac{(66 - 80)^2}{80} = 12,25$$

$$\chi^2_{(0.95, 19)} = 30.14$$

وحيث أن قيمة  $\chi^2$  النظرية (من جدول  $\chi^2$  عند درجات حرية 19 ومساحة 0.95) أكبر من قيمة  $\chi^2$  المحسوبة وفق العلاقة السابقة، فإننا نقبل الفرضية القائلة بأنه لا توجد فروق معنوية بين النسبتين.

## 6.2- اختبار الفرق بين نسبتي:

إذا كان لدينا النسبتين  $\bar{P}_1$  و  $\bar{P}_2$  مأخوذتين من عينتين حجمهما على التوالي:  $n_1$  و  $n_2$ ، من مجتمعين موزعان طبيعيا نسبتهما على التوالي:  $P_1$  و  $P_2$ ، فإنه يمكن اختبار الفرضيتين التاليتين:

$$\begin{aligned} H_0: P_1 &= P_2 \\ H_1: P_1 &> P_2 \end{aligned}$$

وتكون دالة الاختبار:

$$Z = \frac{\bar{P}_1 - \bar{P}_2}{\sigma_{P_1 - P_2}} \quad (2.12)$$

$$\sigma_{P_1 - P_2} = \sqrt{\frac{\bar{p}(1 - \bar{p})}{n_1} + \frac{\bar{p}(1 - \bar{p})}{n_2}} \quad (2.13) \quad \text{حيث:}$$

$$\bar{p} = \frac{n_1 \bar{P}_1 + n_2 \bar{P}_2}{n_1 + n_2} \quad (2.14) \quad \text{و}$$

➤ مثال:

ترغب شركة في أن تحدد عند مستوى معنوية 0.01 في معرفة إذا ما كان نسبة القبول من المكونات الإلكترونية لمورد أجنبي  $P_1$  تزيد عنها لمورد محلي  $P_2$ . أخذت الشركة عينة عشوائية من شحنة كل مورد ووجدت النسب التالية: 0.9 للمورد الأول، بحجم عينة 100 وحدة و 0.7 للمورد الثاني بحجم عينة 80 وحدة. اختبر إذا ما كانت النسب متساوية.

➤ الحل:

$$\begin{aligned} H_0: P_1 &= P_2 \\ H_1: P_1 &> P_2 \end{aligned}$$

بالمقابل يكون:

$$Z = \frac{\bar{P}_1 - \bar{P}_2}{\sigma_{P_1 - P_2}} = \frac{0.2}{0.06} = 3.33$$

حيث أن:

$$\bar{p} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2} = \frac{100.0.9 + 80.0.7}{180} = 0.8$$

$$\sigma_{p1 - p2} = \sqrt{\frac{\bar{p}(1 - \bar{p})}{n_1} + \frac{\bar{p}(1 - \bar{p})}{n_2}} = \sqrt{\frac{0.8.0.2}{100} + \frac{0.8.0.2}{80}} = 0.06$$

وحيث أن قيمة Z النظرية تساوي 2.57 أقل من تلك المحسوبة فإننا نرفض  $H_0$  لصالح  $H_1$ ، أي أن نسبة المورد الأول أكبر معنويا من نسبة المورد الثاني.

## 7.2- اختبارات الكاي مربع $\chi^2$

### 1.7.2- مقارنة بين التكرارات المشاهدة مع التكرارات المتوقعة (اختبارات المطابقة):

تستخدم صيغة الـ  $\chi^2$  في معرفة إذا ما كانت التكرارات المشاهدة الناجمة عن تجربة ما، تطابق التكرارات النظرية (المتوقعة) لهذه التجربة. تكون هذه الصيغة على النحو التالي:

$$\chi^2 = \frac{\sum (f_o - f_e)^2}{f_e} \quad (2.15)$$

حيث:  $f_o$  هي التكرارات المشاهدة و  $f_e$  هي التكرارات النظرية. يجرى الإختبار كما يلي:

هناك تطابق بين التكرارات المشاهدة و التكرارات النظرية:  $H_0$

لا يوجد تطابق بين التكرارات المشاهدة و التكرارات النظرية:  $H_1$

فإذا كانت  $\chi^2$  التطبيقية أكبر من  $\chi^2$  النظرية (التي تحسب من جدول  $\chi^2$  عند مستوى دلالة  $\alpha$  و درجات حرية  $v$ ). تعطى صيغة درجات الحرية بالصيغة التالية:

$$v = k - m - 1$$

k: عدد الفئات و m هي عدد المعالم الإحصائية (المتوسط الحسابي والانحراف المعياري) المقدرة في الاختبار.

➤ مثال:

وجد محل تجاري من خبرته الماضية أن 30 % من السلعة المباعة هي من الحجم الصغير، 40 % هي من الحجم المتوسط و 30 % هي من الحجم الصغير. لتحديد حجم المنتج الواجب الاحتفاظ به من كل نوع، أخذنا عينة حجمها 100 من المبيعات المقدمة، فوجد أن 20 وحدة هي من النوع الكبير، 40 وحدة هي من النوع المتوسط و 40 وحدة هي من النوع الكبير. إختبر إذا ما كان نمط المبيعات الماضي ما زال سائد ( $\alpha = 0.05$ ).

➤ الحل:

لدينا المعطيات التالية:

النمط المشاهد  $f_0$ : 20 40 40

النمط الماضي (المتوقع)  $f_e$ : 30 40 40

$$\chi^2 = \frac{\sum (f_0 - f_e)^2}{f_e} = \frac{(20 - 30)^2}{30} + \frac{(40 - 40)^2}{40} + \frac{(40 - 30)^2}{30} = 5.83$$

وحيث أن:  $v = k - m - 1 = 3 - 0 - 1 = 2$

$m = 0$  لأنه لم يتم حساب أي من معالم المجتمع.

من جهة ثانية نحسب  $\chi^2$  النظرية من الجدول عند درجات حرية 2 ومساحة 0.95 فنجد:

$$\chi^2 = 5.99 \text{ النظرية}$$

وحيث أن  $\chi^2$  النظرية أكبر من  $\chi^2$  التطبيقية فإننا نقبل  $H_0$  أي أن نمط المبيعات الماضي ما زال سائدا.

تستخدم اختبارات الـ  $\chi^2$  كذلك في معرفة طبيعة التوزيع الإحصائي الذي أخذت منه العينة، هل البيانات المدروسة تتبع التوزيع الطبيعي أم توزيع ذو الحدين أم أي توزيع آخر. و لدراسة ذلك نأخذ المثال التطبيقي التالي:

➤ مثال:

قامت وزارة الصحة و السكان بقيام مسح سكاني للعائلات التي لديها 05 أطفال، وقد تم تسجيل الذكور منهم. يمثل الجدول التالي جزء من هذا المسح مجرى على 320 عائلة:

- هل هذه البيانات تخضع لتوزيع ذو الحدين؟.

عدد الأطفال	عدد العائلات
0	8
1	40
2	88
3	110
4	56
5	18
	المجموع : 320

➤ الحل:

لدينا:

البيانات تخضع لتوزيع ذو الحدين  $H_0$ :  
البيانات لا تخضع لتوزيع ذو الحدين  $H_1$ :

نقوم بحساب احتمالات ذو الحدين وفق العلاقة التالية:

$$P(x = k) = C_n^k p^k q^{n-k}$$

$$P(x = 0) = C_5^0 (0.5)^0 (0.5)^5 = 0.0312$$

$$P(x = 1) = C_5^1 (0.5)^1 (0.5)^4 = 0.1562$$

$$P(x = 2) = C_5^2 (0.5)^2 (0.5)^3 = 0.3125$$

$$P(x = 3) = C_5^3 (0.5)^3 (0.5)^2 = 0.3125$$

$$P(x = 4) = C_5^4 (0.5)^4 (0.5)^1 = 0.1562$$

$$P(x = 5) = C_5^5 (0.5)^5 (0.5)^0 = 0.0312$$

نبحث هن التكرارات المشاهدة و التكرارات المتوقعة و نلخص النتائج في

الجدول التالي:

التكرارات المتوقعة (np) : fe	التكرارات المشاهدة fo
$320 \times 0.0312 = 10$	8
$320 \times 0.1562 = 50$	40
$320 \times 0.3125 = 100$	88
$320 \times 0.3125 = 100$	110
$320 \times 0.1562 = 50$	56
$320 \times 0.0312 = 10$	18

تكون صيغة الـ  $\chi^2$  كما يلي:

$$\chi^2 = \frac{\sum (fo - fe)^2}{fe} = \frac{(8 - 10)^2}{10} + \frac{(40 - 50)^2}{50} + \frac{(88 - 100)^2}{100} + \frac{(110 - 100)^2}{100} + \frac{(56 - 50)^2}{50} + \frac{(18 - 10)^2}{10} = 11.96$$

بالمقابل فإن قيمة الـ  $\chi^2$  النظرية يمكن قراءة من الجدول عند المساحة 0.95 (لأن  $\alpha = 5\%$ ) وعند درجات حرية:

$$v = k - m - 1 = 6 - 0 - 1 = 5$$

فنجد:  $\chi^2 = 11.07$  النظرية

وحيث أن  $\chi^2$  النظرية أقل من الـ  $\chi^2$  المغسوبة فإننا نرفض الفرضية القائلة بأن هذه البيانات تتبع توزيع ذو الحدين.

• ملاحظة:

عند إجراء إختبارات المطابقة، يجب التأكد من أن التكرارات المشاهدة  $f_0$  أكبر من أو يساوي 5. فإذا لم يكن كذلك فإنه يجب دمج الفئات التي تكراراتها أقل من 5 مع الفئة اللاحقة أو السابقة لها.

➤ مثال:

هل البيانات التالية تخضع للتوزيع الطبيعي؟.

حدود الفئات	التكرارات
0.61-1.20	1
1.21-1.80	3
1.81-2.40	4
2.41-3.00	65
3.01-3.60	180
3.61-4.20	328
4.21-4.80	408
4.81-5.40	284
5.41-6.00	83
6.01-6.60	13
6.61-7.20	1
7.21-7.80	1
7.81-8.40	0
8.41-9.00	1
	المجموع: 1372

نلاحظ أن تكرارات بعض الفئات أقل من 5، لذلك يجب دمج هذه الفئات. نتائج هذا الدمج وحساب الإحتمالات المقابلة باستعمال التوزيع الطبيعي مسجلة في الجدول الموالي.

إن حساب الإحتمالات المقابلة تتطلب حساب المتوسط الحسابي والانحراف المعياري لهذه البيانات، فنجد:

$$\bar{X} = \frac{\sum X_i \cdot f_i}{\sum f_i} = 4.32 \quad , \quad \sigma = s = \sqrt{\frac{\sum (X_i - \bar{X})^2 \cdot f_i}{\sum f_i}} = 0.81$$

حدود الفئات	التكرارات	الإحتمالات	التكرارات المتوقعة
أقل من أو يساوي 2.40	8	0.091	0.091x1372=12.5
2.41-3.00	65	0.0435	59.68
3.01-3.60	180	0.1368	187.96
3.61-4.20	328	0.2549	349.72
4.21-4.80	408	0.2814	386.08
4.81-5.40	284	0.1858	254.92
5.41-6.00	83	0.0702	96.31
أكبر من أو يساوي 6.01	16	0.0183	25.11

ويمكن حساب الإحتمالات المقابلة وفق العلاقة التالية:

$$Z = \frac{X - \mu}{\sigma}$$

فيكون:

$$P(Z \leq \frac{2.41 - 4.32}{0.81}) = 0.091$$

$$P(\frac{2.41 - 4.32}{0.81} \leq Z \leq \frac{3.00 - 4.32}{0.81}) = 0.0435$$

$$P(\frac{3.01 - 4.32}{0.81} \leq Z \leq \frac{3.60 - 4.32}{0.81}) = 0.1368$$



وهكذا.....

نقوم بحساب قيمة الـ  $\chi^2$  وفق العلاقة التالية:

$$\chi^2 = \frac{\sum (f_o - f_e)^2}{f_e} = \frac{(8 - 1250)^2}{1250} + \frac{(65 - 59.68)^2}{59.68} + \dots = 1346$$

نلاحظ أن:

$$v = k - m - 1 = 8 - 2 - 1 = 5$$

$$\alpha = 0.05 \text{ و}$$

$$\chi^2_{\text{النظرية}} = 11.1 \text{ فيكون:}$$

وحيث أن  $\chi^2$  النظرية أقل من الـ  $\chi^2$  المحسوبة فإننا نرفض الفرضية القائلة بأن هذه البيانات تتبع التوزيع الطبيعي.

• تصحيح ياتس Yates:

تستعمل صيغة كاي مربع لإختبار البيانات المتصلة و المنفصلة، فإذا كانت البيانات منفصلة (متقطعة) فإنه يستحسن تحويلها إلى بيانات متصلة من خلال صيغة ياتس المصححة وفق العلاقة التالية:

$$\chi^2_{\text{المصححة}} = \frac{\sum (|f_o - f_e| - 0.5)^2}{f_e} \quad (2.16)$$

وعموما فإن هذه الصيغة تستخدم عندما تكون درجات الحرية تساوي 1. أما في حالة العينات الكبيرة فإننا نحصل على نتائج مماثلة بين صيغة الكاي مربع العادية و الصيغة المصححة، ولكن قد تنجم بعض الصعوبات بالقرب من القيم الحرجة، لذلك يستحسن إجراء مقارنة بين القيمتين، فإذا كان الفرق كبيرا فإننا نلجأ إلى الرفع من حجم العينة.

➤ مثال:

ترمى قطعة نقود 200 مرة، سجلنا 115 مرة ظهور صورة و 85 مرة ظهور كتابة. إختبر إذا ما كانت القطعة غير متزنة عند  $\alpha = 5\%$ .

➤ الحل:

إن هذا الإختبار يخضع لتوزيع ذو الحدين، فإذا كانت القطعة متزنة فإننا نتحصل على:

$$P = 0.5, q = 0.5, \mu = np = 200 \times 0.5 = 100$$

إذن التكرارات المتوقعة:

$$fe_1 = fe_2 = 100$$

صيغة الـ  $\chi^2$  المحسوبة:

$$\chi^2 = \frac{\sum (fo - fe)^2}{fe} = \frac{(115 - 100)^2}{100} + \frac{(85 - 100)^2}{100} = 4.50$$

نلاحظ أن:

$$v = k - m - 1 = 2 - 1 = 1$$

$$\alpha = 0.05 \text{ و}$$

$$\chi^2_{\text{النظرية}} = 3.84 \text{ فيكون:}$$

وحيث أن  $\chi^2$  المحسوبة أكبر من  $\chi^2$  النظرية فإننا نرفض  $H_0$ ، أي أن القطعة غير متزنة.

$$v = 1$$

وحيث أن:

فإننا نبحث عن قيمة  $\chi^2$  المصححة، فيكون:

$$\chi^2 = \frac{\sum (|fo - fe| - 0.5)^2}{fe} = \frac{\sum (|115 - 100| - 0.5)^2}{100} + \frac{\sum (|85 - 100| - 0.5)^2}{100} = 4.20$$

تبقى  $\chi^2$  المصححة أكبر من  $\chi^2$  النظرية وبالتالي نؤيد القرار السابق، أي أن القطعة غير متزنة.

## 2.7.2- اختبارات كاي مربع للإستقلال و التجانس:

إن في كثير من الحالات نقوم بمعالجة مجموعة من المشاهدات وفق أسلوبين، أي وفق متغيرين  $X$  و  $Y$ ، بما يعرف بالجداول المتصالبة. من خلال اختبارات الكاي مربع يمكننا معرفة إذا ما كان هناك إستقلال أو إرتباط بين المتغيرين  $X$  و  $Y$ .

يقوم مبدأ اختبار الكاي مربع على إيجاد مقياس للأخطاء الناجمة عن تقريب القيم المشاهدة بالقيم المتوقعة بفرض الإستقلال، حيث أن:

$$\chi^2 = \frac{\sum_{i,j} (fo_{ij} - fe_{ij})^2}{fe_{ij}} \quad (2.17)$$

حيث أن:

$fo_{ij}$ : عدد المشاهدات المشاهدة الواقعة في السطر الذي ترتيبه  $i$  و العمود الذي ترتيبه  $j$

$fe_{ij}$ : عدد المشاهدات المتوقعة الواقعة في السطر الذي ترتيبه  $i$  و العمود الذي ترتيبه  $j$

أما قيمة المشاهدات المتوقعة فهي:

$$fe_{ij} = \frac{n_{i+} \cdot n_{j+}}{n} \quad (2.18)$$

$n_{i+}$ : مجموع أعداد المشاهدات الواقعة في السطر الذي ترتيبه  $i$

$n_{j+}$ : مجموع أعداد المشاهدات الواقعة في العمود الذي ترتيبه  $j$

$n$ : عدد المشاهدات الكلي.

أما درجات الحرية فهي:

$$v = (1 - \text{عدد الأسطر})(1 - \text{عدد الأعمدة})$$

## ➤ مثال:

- نريد معرفة تأثير التبغ على القدرة الدراسية لـ 110 طالب. وقد دونت النتائج في الجدول المتصالب التالي:
- إختبر إذا ما كان هناك تأثير للتبغ على القدرة الدراسية للطلاب ( $\alpha=0.05$ )

المجموع	ضعيف	متوسط	ممتاز	مستوى الطالب
				شدة إستهلاك التبغ
30	5	18	7	بكثرة
30	15	7	8	متوسط
50	2	41	7	لا يدخن
110	22	66	22	المجموع

## ➤ الحل:

نقوم بوضع الفرضيتين التاليتين:

$H_0$ : X و Y مستقلان

$H_1$ : X و Y غير مستقلين

المجموع	ضعيف	متوسط	ممتاز	مستوى الطالب
				شدة إستهلاك التبغ
30	5 6	18 18	7 $\frac{22 \times 30}{110} = 6$	بكثرة
30	15 6	7 18	8 6	متوسط
50	2 10	41 30	7 10	لا يدخن
110	22	66	22	المجموع

نقوم بحساب التكرارات المتوقعة وفق العلاقة (2.18):

$$fe_{11} = \frac{n_{i+} \cdot n_{j+}}{n} = \frac{22 \times 30}{110} = 6$$
$$fe_{12} = \frac{n_{i+} \cdot n_{j+}}{n} = \frac{66 \times 30}{110} = 18$$
$$fe_{13} = \frac{n_{i+} \cdot n_{j+}}{n} = \frac{22 \times 30}{110} = 6$$

وهكذا لباقي القيم.....  
نقوم بحساب  $\chi^2$  المحسوبة فنجد:

$$\chi^2 = \sum_{ij} \frac{(fo_{ij} - fe_{ij})^2}{fe_{ij}} = \frac{(7-6)^2}{6} + \frac{(18-18)^2}{18} + \frac{(2-10)^2}{10} = 32.55$$

أما درجات الحرية فهي:

$$v = (عدد الأعمدة - 1)(عدد الأسطر - 1) = (3 - 1)(3 - 1) = 4$$
$$\chi^2_{النظرية} = 9.48 (\alpha = 5\%)$$

وحيث أن  $\chi^2$  المحسوبة أكبر من  $\chi^2$  النظرية فإننا نرفض  $H_0$ ، أي أن هناك  
إستقلال بين الظاهرتين وبالتالي لا توجد علاقة بين إستهلاك التبغ و القدرة  
الدراسية للطلاب.

## تمارين مختارة

### التمرين الأول:

يرغب منتج أسلاك كهربائية إختبار إذا ما كانت هذه الأسلاك لديها قوة مقاومة للكسر تساوي 500 و.د؛ فقوة مقاومة أقل لن تكون ملائمة، و قوة مقاومة أكبر سترفع التكاليف. يأخذ المنتج عينة عشوائية حجمها 64 وحدة ووجد أن لديها قوة مقاومة 510 و.د بإنحراف معياري 48. هل يجب أن يقبل المنتج الفرضية بأن لأسلاكه قوة مقومة 500 و.د عند مستوى دلالة 5%.

### التمرين الثاني:

يرغب مشتري أن يقرر عند مستوى دلالة 5%، أي صنف يشتري من صنفين لهما نفس السعر من مصابيح كهربائية. أخذ المشتري عينة عشوائية حجمها 100 مصباح من كل صنف، ووجد أن مصابيح الصنف الأول تعيش في المتوسط 980 ساعة، بإنحراف معياري 80 ساعة، بينما تعيش مصابيح الصنف الثاني في المتوسط 101 ساعة، بإنحراف معياري 120 ساعة. أي الصنفين من المصابيح يجب شراؤه إذا كان المشتري يرغب في الوصول إلى قرار عند 5% و 1%.

### التمرين الثالث:

يمثل الجدول التالي توزيع أفراد عينتين من مجتمع إحصائي مصنفين حسب فئات السن و حسب نشاطهم الإقتصادي (بطالين، شغيلين)

العينة B		العينة A		فئات السن
شغيلون	بطالون	شغيلون	بطالون	
500	20	100	5	20-30
400	25	100	5	30-40
200	30	300	20	40-50
300	35	500	70	50-60
100	40	500	200	60-70

- هل متوسط السن لفئة البطالين يختلف معنويا في العينتين؟،
- هل توجد فروق معنوية للنسب للفئة الشغيلة؟.

#### التمرين الرابع:

لنفترض أن 50% من 60 مصنع في المنطقة A تخضع لمراقبة معايير التلوث، في حين أن 40% فقط من 40 مصنع موجودة في إقليم B تخضع لنفس العملية. هل نسبة المصانع التي تخضع لمراقبة معايير التلوث أكبر معنويا في A عنها في إقليم B؟. نخذ  $\alpha = 5\%$

#### التمرين الخامس:

يمثل الجدول المقابل توزيع عدد مرات قبول 100 طالب من قبل ثلاث جامعات مختلفة. فإذا كان احتمال قبول طالب ما في جامعة هو 0.4 ؛ إختبر إذا ما كان هذا التوزيع يخضع لذو الحدين ( $\alpha = 5\%$ )

عدد مرات القبول	عدد الطلبة
0	25
1	34
2	31
3	10

#### التمرين السادس:

هل البيانات المعطاة في الجدول المقابل تخضع للتوزيع الطبيعي؟.

التكرارات	حدود الفئات
9	0.655-0.965
13	0.965-1.275
15	1.275-1.585
23	1.585-1.895
21	1.895-2.205
10	2.205-2.515
9	2.515 وأكثر

### التمرين السابع:

في دراسة إحصائية لنتائج إمتحان مسابقة ماء، نريد معرفة إذا ما كانت هناك علاقة بين عامل نجاح الطالب وبين مهنة أب الطالب ( $\alpha = 5\%$ ). نتائج هذه الدراسة مدونة في الجدول التالي:

عدد الطلبة الناجحين	عدد الطلبة الممتحنين	مهنة الأباء
180	2244	موظفون
89	988	سلك التعليم
48	575	أعمال حرة
37	423	بنوك
13	287	تجار وحرفيون
17	210	فلاحة
17	209	مؤسسات مصغرة
المجموع: 402	المجموع: 4936	

- هل يوجد تأثير لمهنة الأباء على قدرة الطالب في النجاح؟.

### التمرين الثامن:

أنتج مجمع صيدال نوعين من مسحوق دواء. لإختبار مفعول هذين النوعين ضد مرض معين، أخذت عينة عشوائية مكونة من 100 شخص مصابة بهذا المرض. بعد تلقيح الأشخاص بالنوعين السابقين، دونت النتائج في الجدول التالي، حيث تعطى إشارة + لحالة نجاح الدواء و إشارة - لعدم نجاحه.



عدد المرضى	النوع A	النوع B
35	-	-
5	+	-
15	-	+
45	+	+

- أعد بناء الجدول،
- هل للنوعين من الدواء تأثير على المرض ( $\alpha = 5\%$ ) .

#### التمرين التاسع:

أرض زراعية مكونة 60 قطعة، قسمت إلى قسمين لكل قسم 30 قطعة. تم تزويد الـ 30 قطعة الأولى بسماد معين، في حين تركت الـ 30 الثانية كشاهد. أعطى إنتاج القسم الأول في المتوسط 18.2 قنطار بـانحراف معياري 0.56 قنطار؛ في حين كان الإنتاج في القسم الثاني 17.5 قنطار بـانحراف معياري 0.43. هل هناك تأثير للسماد على المنتج الزراعي ( $\alpha = 5\%$ )؟



## الفصل الثالث

### نظرية العينات

#### مقدمة:

تعتمد دراسة المجتمعات أساساً على أخذ كل مفردات المجتمع للتعرف على خصائصه، وكيفية توزيع هذه المفردات وحساب معامله. ويقصد بالمجتمع، في الدراسات الإحصائية، مجموعة كل المفردات التي تتصف بوحدة أو أكثر من الصفات المميزة و المشتركة. ويمكن أن نميز طريقتين لدراسة مجتمع ما:

أ- طريقة المسح الشامل، حيث تجمع بيانات المجتمع كله دون إستثناء، مثل تعداد السكان،

ب- طريقة العينة، حيث توجد حالات كثيرة يتعذر فيها إستعمال طريقة المسح الشامل، فيلجأ الإحصائي إلى إنتقاء مجموعة من العينات، قصد إنجاز الدراسة، وتطبق نتائج العينة على بيانات المجتمع.

#### 1.3- طرق إختيار العينة:

إن إختيار العينة من تالمجتمع الإحصائي، يخضاً إلى قاعدتين:

– تحديد هدف الدراسة الإحصائية،

– تعميم نتائج العينة على المجتمع الإحصائي.

#### 1.1.3- طريقة العينة العشوائية البسيطة:

تتميز بكونها مجموعة جزئية من المجتمع الإحصائي، وبحجم معين ولها نفس فرصة الإنتقاء (نفس إحتمال الإختيار). ويمكن إختيار العينة العشوائية البسيطة بالطرق التالية:

– طريقة السحب بدون إرجاء، أي رفض أي عينة أخذت في قراءة سابقة. إن عدد العينات الممكن تشكيلها في هذه الحالة:

$$C_N^n = \frac{N!}{n!(N-n)!} \quad (3.1)$$

– طريقة السحب مع الإرجاء، في هذه الحالة يمكن إختيار عنصر تمت قراءته في سحب سابق. ويكون عدد العينات الممكن تشكيلها في هذه الحالة:

$$C_{N+n-1}^n = \frac{(N+n-1)!}{n!(N-1)!} \quad (3.2)$$

### طريقة الأرقام العشوائية:

وهي إحدى الأساليب في الحصول على عينة عشوائية، وذلك بإعطاء رقم لكل مفردة من المجتمع ومن ثم تسحب هذه الأرقام لتعالج بعد ذلك من خلال جداول تسمى جداول الأرقام العشوائية والتي صممت خصيصا لهذه الطريقة.

### 2.1.3- طريقة العينة الطبقية:

تستخدم هذه الطريقة لزيادة دقة النتائج، وتعتمد على تقسيم المجتمع الإحصائي إلى مجموعات جزئية، تسمى كل مجموعة بالطبقة، ومن كل طبقة يمكن إختيار عينة عشوائية بسيطة.

### 3.1.3- طريقة العينة العنقودية:

وفيها يقسم المجتمع الإحصائي إلى مجموعات جزئية متصلة بينها في شكل عنقود، ومن كل مجموعة نختار عينة عشوائية بسيطة. مثل تقسيم الولاية إلى دوائر والدوائر إلى بلديات والبلديات إلى أحياء والأحياء إلى عمارات والعمارات إلى سكنات.

### 4.1.3- طريقة العينة المعيارية:

وتتميز بثبات المتوسط الحسابي، الوسيط والانحراف المعياري. مثل قياس نسبة نجاح عملية جراحية أجريت على 80 شخص ثم على 60 شخص ثم على 40 شخص. نلاحظ ثبات المقياس بازدياد حجم العينة.

### 2.3- توزيعات المعاينة للعينة العشوائية البسيطة:

إن الهدف من أية عملية معاينة هي تقدير المعالم الأساسية للمجتمع الإحصائي الكلي (المتوسط الحسابي، الانحراف المعياري...) عن طريق معالم العينة، ومن ثم حساب المعالم المقابلة لها في المجتمع.

### 1.2.3- توزيعات المعاينة للأوساط الحسابية:

إذا كانت لدينا مجموعة من العينات مأخوذة من مجتمع ما وسطه الحسابي  $\mu$  وإنحرافه المعياري  $\sigma$ ، فإن معظم الأوساط الحسابية لهذه العينات تختلف عن بعضها. نسمي التوزيع الإحصائي لمتوسطات هذه العينات بتوزيع المعاينة للوسط الحسابي، ولسوف نجد أن لهذه الأوساط (التوزيع) متوسط حسابي يرمز له بالرمز  $\mu_{\bar{x}}$ ، حيث أن:

$$\mu_{\bar{x}} = \frac{\bar{X}_1 + \bar{X}_2 + \dots + \bar{X}_n}{C_N^n} = \mu \quad (3.3)$$

أما إنحرافه المعياري:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}} \sqrt{\frac{N-n}{N-1}} \quad (3.4) \text{ في حالة مجتمع محدود (سحب بدون إرجاع)}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}} \quad (3.5) \text{ في حالة مجتمع غير محدود (سحب مع الإرجاع).}$$

ولقيم  $n$  الكبيرة ( $n > 30$ ) فإن توزيع المعاينة للأوساط الحسابية تقترب من التوزيع الطبيعي بمتوسط  $\mu_{\bar{x}}$  وبانحراف معياري  $\sigma_{\bar{x}}$ . إن هذه الخاصية في الحقيقة تنطبق على المجتمعات الغير المحدودة والتي تثبت دقة التقريب من التوزيع الطبيعي، وهي حالة خاصة من نظرية النهاية المركزية، التي سترى لاحقاً.

➤ مثال:

مجتمع إحصائي حجمه  $N = 05$  يمثل أعمار 05 أطفال هي:  
 $X_1=6, X_2=8, X_3=10, X_4=12, X_5=14$

لنقسم المجتمع إلى مجموعة من العينات، حجم كل عينة 2 وذلك بالسحب بدون إرجاع. إن عدد العينات اللازم تشكيلها في هذه الحالة هي:

$$\mu = \frac{6+8+10+12+14}{5} = 10 \quad \text{وأن:} \quad C_5^2 = \frac{5!}{2!(5-2)!} = 10$$

$\sigma = 2.82$       كما أن:

الوسيط الحسابي للعينة	العينة
7	(6,8)
8	(6,10)
9	(6,12)
10	(6,14)
9	(8,10)
10	(8,12)
11	(8,14)
11	(10,12)
12	(10,14)
13	(14,12)

نجد:

$$\mu_{\bar{x}} = \frac{7+8+9+...+12+13}{10} = 10 = \mu$$

أما الانحراف المعياري في حالة مجتمع محدود:

$$\sigma_{\bar{x}} = \frac{2.82}{\sqrt{5}} \sqrt{\frac{5-2}{5-1}} = 1.73 \neq \sigma$$

### • حالة عينتين من مجتمعين مستقلين:

إذا كانت  $X_1, X_2, \dots, X_n$  عينة مأخوذة من توزيع وسطه الحسابي  $\mu_1$  وانحرافه المعياري  $\sigma_1$ ، وكانت  $Y_1, Y_2, \dots, Y_n$  عينة أخرى من توزيع آخر وسطه الحسابي  $\mu_2$  وانحرافه المعياري  $\sigma_2$ ، فإنه يمكن الحصول على توزيع الفروق للأوساط الحسابية وفق العلاقة التالية:

$$\mu_{\bar{x}} - \mu_{\bar{y}} = \mu_1 - \mu_2$$

أما الانحراف المعياري لهذا الفرق فيعرف بالعلاقة التالية:  
في حالة مجتمع غير محدود:

$$\sigma_{\bar{x} - \bar{y}} = \sqrt{\sigma_{\bar{x}}^2 + \sigma_{\bar{y}}^2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad (3.6)$$

في حالة مجتمع محدود:

$$\sigma_{\bar{x} - \bar{y}} = \sqrt{\sigma_{\bar{x}}^2 + \sigma_{\bar{y}}^2} = \sqrt{\frac{\sigma_1^2}{n_1} \cdot \frac{N_1 - n_1}{N_1 - 1} + \frac{\sigma_2^2}{n_2} \cdot \frac{N_2 - n_2}{N_2 - 1}} \quad (3.7)$$

### ➤ مثال:

أخذت عينة حجمها 30 وحدة من توزيع وسطه الحسابي 75 و تباينه 25، وأخذت عينة ثانية حجمها 60 من توزيع آخر مستقل وسطه الحسابي 60 وتباينه 15. أوجد الفروق للوسط الحسابي والانحراف المعياري لتوزيع المعاينة للتوزيعين.

### ➤ الحل:

بما أن المجتمعين غير محدودين فإن:

$$\mu_{\bar{x}} - \mu_{\bar{y}} = 75 - 60 = 15$$

أما الفروق بين التباينين:

$$\sigma_{\bar{x} - \bar{y}} = \sqrt{\frac{25}{30} + \frac{15}{60}} = 1.04$$

### • نظرية النهاية المركزية:

إذا كان حجم العينة كبيرا ( $n > 30$ ) فإن توزيع المعاينة للمتوسط الحسابي يقترب من التوزيع الطبيعي، بغض النظر عن شكل المجتمع الأصلي. لذلك يمكن حساب احتمال أن يكون  $\bar{X}$  لعينة عشوائية داخل فترة معينة من خلال حساب القيمة الإحصائية لـ  $Z$  وفق العلاقة الشهيرة التالية:

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} \quad (3.8)$$

➤ مثال:

صممت إحدى الشركات سيارة بحيث أن أكبر حمولة لها هي 3000 كغ وتتسع إلى 30 راكبا. إذا علمت أن الأوزان تخضع للتوزيع الطبيعي الذي وسطه 70 كغ وانحرافه المعياري 54.77. أحسب احتمال أن تحمل هذه السيارة أكثر من حمولتها.

➤ الحل:

$$\text{نلاحظ أن: } \bar{X} = \frac{3000}{30} = 100 \quad \text{والمطلوب إيجاد: } P(\bar{X} > 100)$$

لدينا:

$$P(\bar{X} > 100) = P\left(\frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} > \frac{100 - 70}{\frac{54.77}{\sqrt{30}}}\right) =$$

$$P(z > 3) = 0.5 - 0.4987 = 0.0013$$

### • حالة مجتمعين مستقلين:

إذا كانت  $X_1, X_2, \dots, X_n$  عينة مأخوذة من توزيع وسطه الحسابي  $\mu_A$  وانحرافه المعياري  $\sigma_A$ ، وكانت  $Y_1, Y_2, \dots, Y_n$  عينة أخرى من توزيع آخر مستقل



وسطه الحسابي  $\mu_B$  وانحرافه المعياري  $\sigma_B$ ، فإننا يمكن إيجاد القيمة الإحصائية  $Z$  وفق العلاقة التالية:

$$Z = \frac{(\bar{X}_A - \bar{X}_B) - \mu_{\bar{X}_A - \bar{Y}_B}}{\sigma_{\bar{X}_A - \bar{Y}_B}} \quad (3.9)$$

➤ مثال:

تنتج الشركة A مصابيح كهربائية متوسط مدة حياتها 1400 ساعة، بانحراف معياري 200 ساعة، وتنتج الشركة B مصابيح مماثلة متوسط مدة حياتها 1200 ساعة بانحراف معياري 100 ساعة. قمنا باختيار عينة عشوائية حجمها 125 وحدة من كل شركة. أوجد احتمال أن الشركة A تنتج مصابيح كهربائية متوسط مدة حياتها على أكبر من الشركة B بـ: - 160 ساعة، 250 ساعة.

➤ الحل:

نلاحظ أن:

$$\begin{aligned} \mu_{\bar{X}_A - \bar{X}_B} &= \mu_{\bar{X}_A} - \mu_{\bar{X}_B} = 1400 - 1200 = 200 \\ \sigma_{\bar{X}_A - \bar{X}_B} &= \sqrt{\sigma_{\bar{X}_A}^2 + \sigma_{\bar{X}_B}^2} = \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} = \sqrt{\frac{200^2}{125} + \frac{100^2}{125}} = 20 \end{aligned}$$

إذن احتمال أن الشركة A تنتج مصابيح كهربائية متوسط مدة حياتها على أكبر من الشركة B بـ 160 ساعة هو:

$$\begin{aligned} P((\bar{X}_A - \bar{X}_B) > 160) &= P\left(\frac{(\bar{X}_A - \bar{X}_B) - \mu_{\bar{X}_A - \bar{Y}_B}}{\sigma_{\bar{X}_A - \bar{Y}_B}} > \frac{160 - 200}{20}\right) \\ &= P(z > -2) = 0.5 + 0.4772 = 0.9772 \end{aligned}$$

وإ احتمال أن الشركة A تنتج مصابيح كهربائية متوسط مدة حياتها على أكبر من الشركة B بـ 250 ساعة هو:

$$\begin{aligned} P((\bar{X}_A - \bar{X}_B) > 250) &= P\left(\frac{(\bar{X}_A - \bar{X}_B) - \mu_{\bar{X}_A - \bar{Y}_B}}{\sigma_{\bar{X}_A - \bar{Y}_B}} > \frac{250 - 200}{20}\right) = \\ P(z > 2.50) &= 0.5 - 0.4938 = 0.0062 \end{aligned}$$

### 2.2.3- المعاينة باستعمال توزيع t:

إذا كان حجم العينة صغيرا ( $n < 30$ ) فإنه يمكن إستخدام توزيع t كبديل للتوزيع الطبيعي. ويمكن حساب إحتمال أن  $\bar{X}$  لعينة عشوائية، داخل فترة معينة، من خلال حساب القيمة الإحصائية لـ t وفق العلاقة:

$$t = \frac{\bar{X} - \mu_{\bar{X}}}{S_{\bar{X}}} \quad \text{بدرجات حرية } n-1 \quad (3.10)$$

➤ مثال:

تخضع علامات طلبة معهد ما للتوزيع الطبيعي بوسط حسابي 70. أخذت عينة عشوائية حجمها 16 طالبا بإنحراف معياري 8. أحسب إحتمال أن تزيد علامات الطلبة عن 78 درجة.

➤ الحل:

المطلوب هو حساب:

$$P(\bar{X} > 74) = P\left(\frac{\bar{X} - \mu_{\bar{X}}}{S_{\bar{X}}} > \frac{74 - 70}{\frac{8}{\sqrt{16}}}\right) = P(t > 2) = 0.025$$

عند درجات حرية 15 نجد المساحة 0.975

### 2.2.3- توزيعات المعاينة للتباينات:

إذا كانت  $X_1, X_2, \dots, X_n$  عينة مأخوذة من توزيع وسطه الحسابي  $\mu$  وتباينه  $\sigma^2$ ، وكان  $s^2$  هو تباين هذه العينة بحيث  $\sigma^2$  معلومة، فإن تباين العينة يخضع لتوزيع كاي مربع وفق العلاقة:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \quad \text{بدرجات حرية } n-1 \quad (3.11)$$

➤ مثال:

إذا كانت  $X_1, X_2, \dots, X_n$  عينة عشوائية من توزيع طبيعي  $N(\mu, 9)$  تباينها  $S^2$ . أوجد النقطة c بحيث يكون:  $P(S^2 \leq c) = 0.95$ .

➤ الحل:

$$P(S^2 \leq c) = P\left(\frac{n-1}{\sigma^2} \cdot S^2 \leq \frac{c(n-1)}{\sigma^2}\right)$$

$$= P\left(\chi^2_{9} \leq \frac{c(10-1)}{9}\right) = 0.95 \Rightarrow P(\chi^2_{9} \leq c) = 0.95$$

وباستعمال جدول الكاي مربع نجد:  $c = 16.91$

• حالة  $\sigma$  مجهولة:

إذا كانت  $X_1, X_2, \dots, X_n$  عينة مأخوذة من توزيع وسطه الحسابي  $\mu$  وتباينه  $\sigma^2$ ، وكان  $s^2$  هو تباين هذه العينة بحيث  $\sigma^2$  مجهولة، فإن تباين العينة يخضع لتوزيع  $t$  وفق العلاقة:

$$t = \frac{X - \mu}{S / \sqrt{n-1}} \quad \text{بدرجات حرية } n-1 \quad (3.12)$$

3.2.3- توزيعات المعاينة للنسب:

لنفترض أنه لدينا مجتمع غير محدود، به جميع العينات الممكنة ذات حجم  $n$ ، ولكل عينة نسبة النجاح  $p$  حيث أن  $p + q = 1$ ، فإن توزيع المعاينة للنسب يكون بمتوسط  $\mu_p$  و إنحراف معياري  $\sigma_p$ ، حيث:

$$\begin{aligned} \mu_p &= p \\ \sigma_p &= \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}} \end{aligned} \quad (3.13)$$

أما إذا كان المجتمع محدود (المعاينة بدون إرجاع) فإن العلاقة تصبح كما

يلي:

$$\begin{aligned} \mu_p &= p \\ \sigma_p &= \sqrt{\frac{pq}{n}} \cdot \sqrt{\frac{N-n}{N-1}} \end{aligned} \quad (3.14)$$

➤ مثال:

ترمى قطعة نقود متزنة 120 مرة. أوجد احتمال الحصول ما بين 40% و60% صورة.

➤ الحل:

إن احتمالات النجاح (الحصول على صورة) هي  $P = \frac{1}{2}$  وإحتمالات الفشل (الكتابة) هي  $q = \frac{1}{2}$  ؛ وحيث أن  $n \geq 30$  فإن احتمالات ذو الحدين تقترب من التوزيع الطبيعي، ويكون:

$$48 = 120 \times 40\% \quad \text{و} \quad 72 = 120 \times 60\%$$

أي المطلوب حساب  $P(48 \leq X \leq 72)$  ؛ و بما أننا نقوم بعملية تقريب توزيع منفصل إلى توزيع متصل فإن الإحتمال السابق يأخذ الشكل التالي:

$$P(47.5 \leq X \leq 72.5)$$

لدينا:

$$\mu_p = np = 120 \cdot \frac{1}{2} = 60$$

$$\sigma = \sqrt{npq} = \sqrt{120 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 5.48$$

ومنه:

$$P(47.5 \leq X \leq 72.5) = P\left(\frac{47.5 - 60}{5.48} \leq Z \leq \frac{72.5 - 60}{5.48}\right) =$$

$$P(-2.28 \leq Z \leq 2.28) = 0.9774$$

## تمارين مختارة

### التمرين الأول:

مجتمع مكون من 12000 عنصرا بوسط حسابي 100 وانحراف معياري 60. أوجد الوسط الحسابي والانحراف المعياري لتوزيع المعاينة للمتوسط الحسابي، عندما يكون حجم العينة 100.

### التمرين الثاني:

مجتمع مكون 3000 امرأة متوسط أعمارهن 25 سنة. فإذا كانت أوزان هؤلاء النسوة تتبع التوزيع الطبيعي بوسط 58 كغ وانحراف معياري 4، أخذت 80 عينة عشوائية حجمها 25 امرأة لكل عينة. أوجد المتوسط الحسابي والانحراف المعياري لتوزيع المعاينة، وذلك في حالة:  
- السحب بدون إرجاع - السحب بالإرجاع

### التمرين الثالث:

إذا كان متوسط وزن 500 كرة هو 5.02 كغ، بانحراف معياري 0.3 كغ. أخذت عينة عشوائية حجمها 100؛ أوجد احتمال أن يكون وزن الكريات:  
- محصورا بين 496 و 500 كغ؛ - أكبر من 510 كغ

### التمرين الرابع:

بينت الإحصائيات في بلد ما أن 64% من السكان يملكون سيارة سياحية. أخذت عينة عشوائية مكونة من 225 شخص؛ أوجد احتمال أن نسبة الذين يملكون سيارة سياحية: - محصورا بين 0% و 70%، - أكبر من 60%، - أقل من 25%.

### التمرين الخامس:

ماكينة صناعية مختصة في إنتاج قطع حديد دائرية متوسط قطرها 25مم. بعد مدة أصاب الماكينة عطل وتم تصليحه؛ أخذت عينة عشوائية من منتج

الماكينة بعد تصليحها وتم قياس الأقطار المنتجة فكانت النتائج كما يلي (بالملم):  
22، 23، 21، 25، 24، 22، 26، 21. هل يمكن الجزم بأن الماكينة مازالت  
تعمل بشكل جيد، عند احتمال 95 % ، بفرض أن أقطار القطع تتبع التوزيع  
الطبيعي؟.

#### التمرين السادس:

أظهرت دراسة طبية وجود نوعين من مرض الروماتيزم، الإلتهابي وغير  
الإلتهابي؛ من أصل 220 مصاب بمرض الروماتيزم، وجد أن 167 شخص  
مصابون بالنوع الإلتهابي.

- أعطي تقديرا نقطيا وتقديرا بمجال لنسبة الأشخاص المصابون بالنوع  
الإلتهابي، قامت هذه الدراسة بتحليل عامل المناعة  $X$  في دم المصابين بالنوعين  
من المرض، فكانت النتائج كما يلي:

	النوع الإلتهابي	النوع الغير الإلتهابي
$\sum X$	420	104
$\sum X^2$	1400	292

- أوجد مجال الثقة لمتوسط عامل المناعة  $X$  ( $\alpha = 10\%$ ) لدى النوعين من  
المرض. ما ذا تقترح؟.

#### التمرين السابع:

ينتج مصنعا للأدوية نوعا من الدواء يحتوي على مادة فعالة، و يجب أن  
تكون هذه المادة محددة بشكل دقيق. و لدراسة هذه الدقة، قمنا بتحليل عينة  
حجمها 25 حبة. أحسب احتمال أن لا يزيد الانحراف المعياري لكمية هذه  
المادة في الحبوب عن 1.30 غ، علما أن الانحراف المعياري لوزن المادة الفعالة في  
إنتاج المصنع كله هو 1.20 غ.

## الفصل الرابع التقدير

### مقدمة:

تعتمد التوزيعات الإحصائية على معالم، فمثلا يعتمد التوزيع الطبيعي على المعلمتين  $\mu$  و  $\sigma$ ، ويعتمد توزيع بواسون على المعلمة  $\lambda$ ، أما توزيع ذو الحدين فيعتمد على  $P$  (إحتمالات النجاح)؛ وعادة ما تكون هذه المعالم مجهولة، وليس من السهل إيجادها، ولذلك نلجأ إلى تقدير هذه المعالم.

### 1.4- معايير جودة التقدير:

#### (أ) - عدم التحيز:

نقول عن المقدار  $\tilde{a}_k$  أنه يعتبر تقديرا غير متحيزا لـ  $a$  فيما إذا كان التوقع الرياضي لـ  $\tilde{a}_k$  يساوي  $a$  أي:

$$E(\tilde{a}_k) = a \quad (4.1)$$

حيث  $k$  يرمز إلى جميع التقديرات الممكنة المأخوذة من جميع العينات الممكنة ذات الحجم  $n$ .  
أما إذا كان:

$$E(\tilde{a}_k) \neq a \quad (4.2)$$

فإن المقدار  $\tilde{a}_k$  أنه يعتبر تقديرا متحيزا لـ  $a$  أي أنه لا يمثل المقدار  $a$  تمثيلا صحيحا.

فمثلا متوسط توزيع المعاينة للأوساط يساوي متوسط المجتمع أي:  $\mu_x = \mu$ ، في حين يكون متوسط العينة  $\bar{X}$  هو تقدير غير متحيز لمتوسط المجتمع.

### ➤ مثال:

قمنا بقياس أقطار شجرة ما وو جدنا النتائج التالية:

6.33, 6.37, 6.36, 6.37, 6.37 (m)

أوجد تقديرات غير متحيزة للمتوسط الحقيقي و التباين الحقيقي.  
➤ **الحل:**

إن التقدير غير المتحيز للمتوسط هو:

$$\bar{X} = \frac{\sum X}{N} = \frac{31.75}{5} = 6.35 \text{ وهو تقديرا غير متحيز لمتوسط المجتمع } \mu$$

أما التباين:

$$S^2 = \frac{\sum (X - \bar{X})^2}{n-1} = 0.000555 \text{ فيكون } S = \sqrt{0.000555} = 0.023 \text{ هو}$$

تقدير متحيز لـ  $\sigma$  لأنه لا يمثل تباين المجتمع تمثيلا صحيحا.

(ب) - التماسك:

نقول عن المقدار  $\tilde{a}_k$  أنه يعتبر تقديرا لـ  $a$  فيما إذا كان:

$$\lim_{n \rightarrow N(\infty)} P(\tilde{a}_k \rightarrow a) = 1 \quad (4.3)$$

حيث  $P$  يرمز إلى الاحتمال، أي أنه كلما كبر حجم العينة فإننا نحصل على أفضل تقدير للمقدار  $a$ ، و هو ما يدعى بالتماسك.

➤ **مثال:**

في حالة توزيعات المعاينة، رأينا أن متوسط متوسط العينات المسحوبة  $\mu_{\bar{x}}$  ما هو إلا متوسط المجتمع  $\mu$  ولذلك فإن هذا المتوسط يعتبر تقديرا متماسكا لمتوسط المجتمع.

(ج) - الفعالية:

إذا كان للمقدار  $a$  عددا كبيرا من التقديرات الفعالة و التماسكة، فإن التقدير الفعال (الذي يجب إختياره) هو ذو التباين الأصغر أي أن:

$$\text{Var}(\tilde{a}_k) = \frac{\sum_{k=1}^M (a_k - a)^2}{M} \rightarrow \text{Min} \quad (4.4)$$



#### 2.4- التقدير النقطي:

وهي أبسط طرق التقدير وتتمثل في تقدير المعلمة بقيمة واحدة، كأن نقول أن متوسط دخل أفراد أسرة ما هو 300 دينار، فيكون  $\bar{X} = 300$  هو تقدير نقطي لمتوسط المجتمع.

#### 3.4- التقدير بمجال:

تنطوي عملية التقدير بنقطة عن بعض الأخطاء، لهذا فإننا نحتاج إلى بعض المعلومات الخاصة عن مدى إنحراف هذا التقدير عن قيمة المعلمة. نكون في هذا الحالة قد كونا مجال معين تتغير فيه قيمة المعلمة تحت احتمال معين.

#### 1.3.4- مجالات الثقة للأوساط الحسابية:

لا حظنا أن  $\bar{X}$  هو تقدير لـ  $\mu$ ، وحيث أن  $\bar{X}$  هو متوسط حسابي لعينة، فإن تقدير  $\mu$  بـ  $\bar{X}$  تنشأ عنه أخطاء.

إذا كانت  $X_1, X_2, \dots, X_n$  عينة عشوائية مأخوذة من توزيع طبيعي  $N(\mu, \sigma^2)$  فإن  $\bar{X}$  يخضع للتوزيع الطبيعي  $N(\mu, \sigma^2/n)$  حيث:

$$P(-2 \leq Z \leq +2) = P(-2 \leq \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq +2) = 0.95$$

$$P(-2 \frac{\sigma}{\sqrt{n}} \leq \bar{X} - \mu \leq +2 \frac{\sigma}{\sqrt{n}}) = P(\mu - 2 \frac{\sigma}{\sqrt{n}} \leq \bar{X} \leq \mu + 2 \frac{\sigma}{\sqrt{n}}) = 0.95$$

فيكون مجال تقدير  $\bar{X}$  هو  $\left[ \mu - 2 \frac{\sigma}{\sqrt{n}}, \mu + 2 \frac{\sigma}{\sqrt{n}} \right]$ ، وحيث أن  $\mu$  هي قيمة مجهولة وأن  $\bar{X}$  هو تقدير نقطي لـ  $\mu$  فإن لـ  $\mu$  هو:

$$\left[ \bar{X} - 2 \frac{\sigma}{\sqrt{n}}, \bar{X} + 2 \frac{\sigma}{\sqrt{n}} \right] \quad (4.5)$$

➤ مثال:

أخذت عينة حجمها 30. بمتوسط حسابي 22، فإذا كانت العينة مأخوذة من توزيع طبيعي  $N(\mu, 45)$ ، فما هو مجال الثقة لمتوسط المجتمع الذي أخذت منه العينة؟.

➤ الحل:

بتطبيق العلاقة (4.5) نجد:

$$\left[ 22 - 2 \frac{\sqrt{45}}{\sqrt{30}}, 22 + 2 \frac{\sqrt{45}}{\sqrt{30}} \right] = [19.55, 24.44]$$

باحتمال قدره 95.45 %

• نظرية:

إذا كانت  $X_1, X_2, \dots, X_n$  عينة عشوائية مأخوذة من توزيع طبيعي  $N(\mu, \sigma^2)$  حيث  $\sigma^2$  معلومة، فإن مجال الثقة  $(1-\alpha)\%$  للمتوسط الحسابي  $\bar{X}$  هو:

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad (4.6)$$

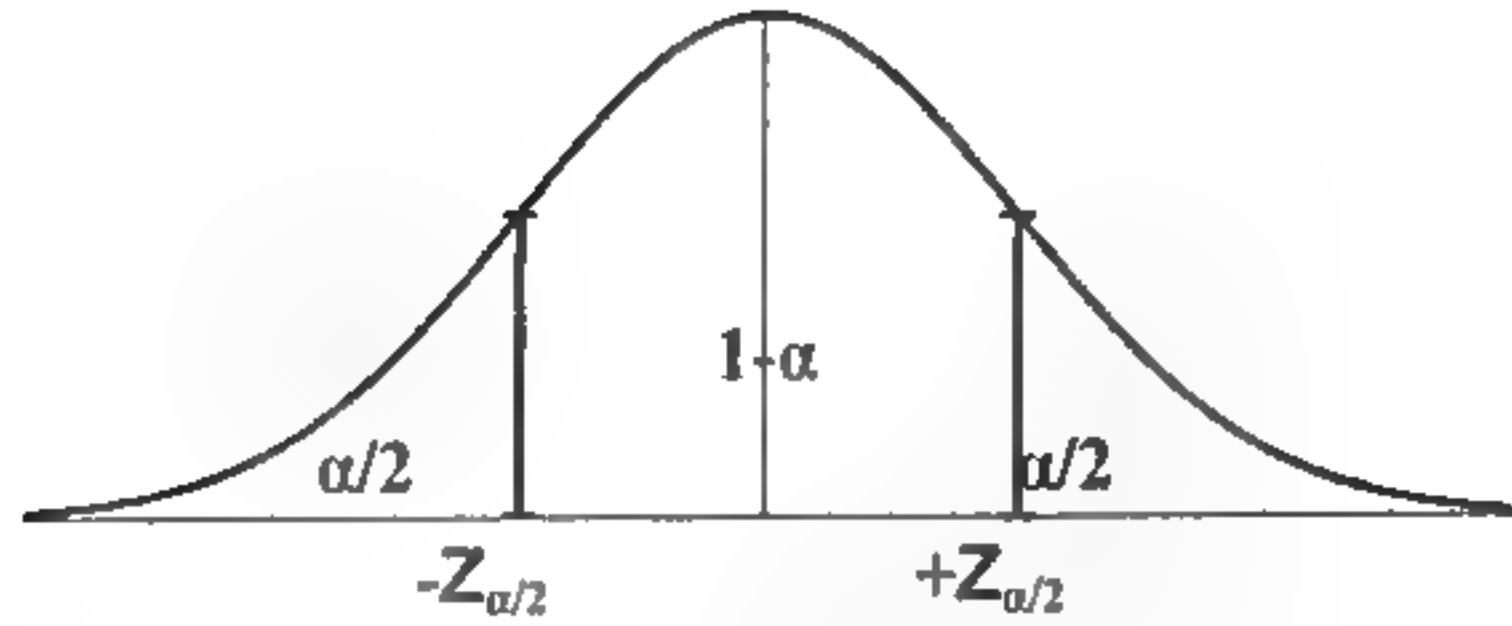
وبصفة عامة يمكن أن نكتب القوانين السابقة كما يلي:

$$P(-Z_{\alpha/2} \leq Z \leq +Z_{\alpha/2}) = P(-Z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq +Z_{\alpha/2}) = 1 - \alpha$$

$$P(\bar{X} - Z_{\alpha/2} \cdot \sigma_{\bar{X}} \leq \mu \leq \bar{X} + Z_{\alpha/2} \cdot \sigma_{\bar{X}}) = 1 - \alpha$$

$$\mu \in [\bar{X} - Z_{\alpha/2} \sigma_{\bar{X}}, \bar{X} + Z_{\alpha/2} \sigma_{\bar{X}}]$$

حيث أن:  $P(Z \leq Z_{\alpha/2}) = 1 - \alpha/2$



فتكون قيم  $Z_{\alpha/2}$  حسب قيم  $\alpha$  هي:

$$\alpha = 1\% \Rightarrow Z_{\alpha/2} = 2.58$$

$$\alpha = 5\% \Rightarrow Z_{\alpha/2} = 1.96$$

$$\alpha = 10\% \Rightarrow Z_{\alpha/2} = 1.65$$

• حالة  $n < 30$  :

ويمكن أن نميز:

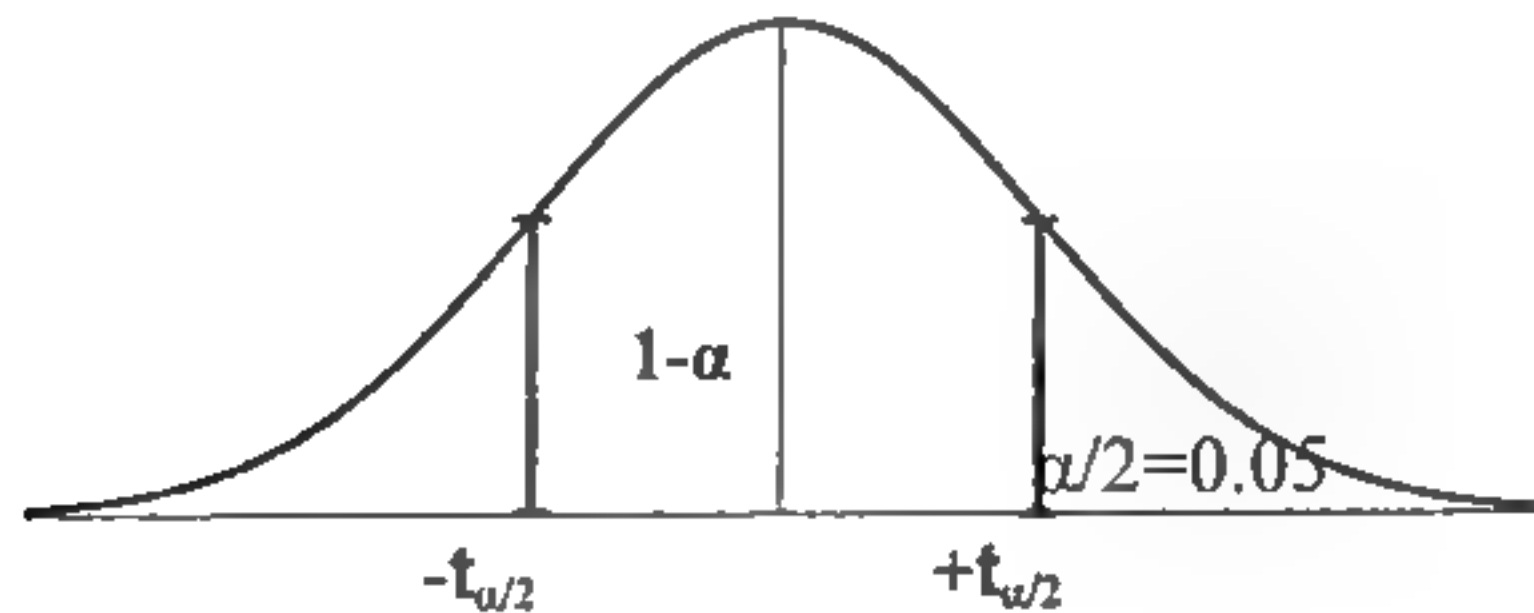
- إذا كانت  $\sigma$  معلومة، فإن مجال الثقة للمتوسط الحسابي هو:

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad (4.7)$$

- إذا كانت  $\sigma$  مجهولة، فإننا نستخدم  $S$  كتقدير لها، و يكون مجال الثقة للمتوسط الحسابي هو:

$$\bar{X} \pm t_{\alpha/2} \frac{S}{\sqrt{n}} \quad (4.8) \quad \text{بدرجات حرية } n-1$$

غير أن العلاقة (4.8) تستخدم حتى في حالة  $n \geq 30$  عندما تكون  $\sigma$  مجهولة.



➤ مثال:

عينة عشوائية حجمها 16 مفردة مأخوذة من  $N(\mu, \sigma^2)$ ، فإذا كان المتوسط الحسابي للعينة هو 14.5 بتباين 25، فأوجد مجال الثقة 90% لمتوسط المجتمع.

➤ الحل:

حيث أن  $n < 30$  فإن مجال الثقة في هذه الحالة:

$$\bar{X} \pm t_{\alpha/2} \frac{S}{\sqrt{n}}$$

$$S_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{5}{4}$$

$$t_{\alpha/2} = t_{0.10/2} = t_{(0.05, 15)} = -t_{(0.95, 15)} = -1.753$$

فيكون:

$$14.5 \pm 1.75 \frac{5}{4} = [12.3, 16.7]$$

$$\mu \in [12.3, 16.7] \quad \text{أي أن:}$$

• تقدير الفرق بين وسطين:

إذا كانت  $\bar{X}$  متوسط حسابي لعينة عشوائية حجمها  $n_1$  مأخوذة من مجتمع 1، وكانت  $\bar{Y}$  متوسط حسابي لعينة عشوائية أخرى حجمها  $n_2$  وكانت مأخوذة من مجتمع 2، حيث أن  $\sigma_1$  و  $\sigma_2$  معلومتين؛ فإن مجال الثقة للفرق بين وسطين هو:

$$(\bar{X} - \bar{Y}) \pm Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad (4.9)$$

أما إذا كانت  $\sigma_1$  و  $\sigma_2$  مجهولتين، فإننا نستخدم، على التوالي  $S_1$  و  $S_2$  كتقدير لهما وذلك بإستعمال توزيع  $t$ ، فيكون:

$$(\bar{X} - \bar{Y}) \pm t_{\alpha/2} S_c \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad (4.10)$$

بدرجات حرية  $n_1+n_2-1$

$$S_c = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad \text{حيث:}$$

#### 2.3.4- تقدير النسب:

إذا كانت  $X_1, X_2, \dots, X_n$  عينة عشوائية مأخوذة من مجتمع يخضع لتوزيع ذو الحدين، وكانت  $p$  هي نسبة النجاح (إحتمالات النجاح)، فإن مجال الثقة للنسبة  $P$  (المجتمع) هي:

$$p \pm Z_{\alpha/2} \sqrt{\frac{pq}{n}} = p \pm Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}} \quad (4.11)$$

حيث:  $P$  المجتمع التي هي مجهولة.

إذا كان حجم المجتمع  $N$  معلوم فإن المعادلة (4.9) تأخذ الشكل التالي:

$$p \pm Z_{\alpha/2} \sqrt{\frac{pq}{n}} \cdot \sqrt{\frac{N-n}{N-1}} \quad (4.12)$$

#### ➤ مثال:

قمنا بمسح إحصائي لتقدير نسبة الذكور في مدينة ما، أخذت عينة عشوائية حجمها 300 شخص ووجد أن عدد الذكور هو 123. أوجد مجال الثقة 95 % لنسبة الذكور في المجتمع.  
لدينا:

$$p = \frac{123}{300} = 0.41$$

ومنه:

$$\bar{p} \pm Z_{\alpha/2} \sqrt{\frac{pq}{n}} = 0.41 \pm 1.96 \sqrt{\frac{0.41 \times 0.59}{300}} = [0.36, 0.46]$$

### 3.3.4- تقدير الفرق بين نسبتي:

بنفس الطريقة المتبعة في تقدير الفرق بين وسطين، يمكن تقدير الفرق بين نسبتي في العلاقة التالية:

$$(p_1 - p_2) \pm Z_{\alpha/2} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \quad (4.13)$$

حيث:

$P_2$ : نسبة العينة الأولى،  $P_2$  هي نسبة العينة الثانية.

### 2.3.4- مجالات الثقة للتباينات:

(أ) - مجال الثقة لتباين المجتمع:

إذا كانت  $X_1, X_2, \dots, X_n$  عينة عشوائية مأخوذة من توزيع طبيعي  $N(\mu, \sigma^2)$ ، فإن مجال الثقة  $(1-\alpha)\%$  لتباين المجتمع:

$$\left[ \frac{(n-1)S^2}{\chi^2_{(1-\alpha/2, n-1)}}, \frac{(n-1)S^2}{\chi^2_{(\alpha/2, n-1)}} \right] \quad (4.14)$$

حيث  $S^2$  هو تباين العينة و  $\chi^2$  هو المتغير العشوائي كاي مربع.

➤ مثال:

أخذنا عينة عشوائية حجمها 10 ووجد أن تباينها يساوي 117.12. أوجد مجال 95% الثقة لتباين المجتمع.

➤ الحل:

لدينا:

$$\begin{aligned} \chi^2_{(1-\alpha/2, n-1)} &= \chi^2_{(0.95, 9)} = 19.02 \\ \chi^2_{(\alpha/2, n-1)} &= \chi^2_{(0.025, 9)} = 2.70 \end{aligned}$$

ومنه، مجال الثقة:

$$\left[ \frac{9 \times 117.12}{19.02}, \frac{9 \times 117.12}{2.70} \right] = [55.41, 390.4]$$

(ب) - مجال الثقة للنسبة بين تباينين:

إذا كانت  $X_1, X_2, \dots, X_n$  عينة عشوائية مأخوذة من توزيع طبيعي  $N(\mu_1, \sigma_1^2)$  وكانت  $Y_1, Y_2, \dots, Y_n$  عينة أخرى عشوائية مأخوذة من توزيع طبيعي  $N(\mu_2, \sigma_2^2)$ ، فإن مجال الثقة  $(1-\alpha)\%$  للنسبة  $\frac{\sigma_1^2}{\sigma_2^2}$  هي:

$$\left[ \frac{S_2^2}{S_1^2} \cdot F_{(\alpha/2, n_1-1, n_2-1)}, \frac{S_2^2}{S_1^2} \cdot F_{(1-\alpha/2, n_1-1, n_2-1)} \right] \quad (4.15)$$

حيث  $S_1^2$  و  $S_2^2$  هما تبايني العينتين على التوالي،  $F$  هو المتغير العشوائي فيشر.

## تمارين مختارة

### التمرين الأول:

أخذت عينة عشوائية مكونة من 25 مفردة بمتوسط 80، وانحراف معياري 30، من مجتمع حجمه 1000 مفردة ويتبع التوزيع الطبيعي. أوجد مجال الثقة 90 % لمتوسط المجتمع.

### التمرين الثاني:

لمعرفة تأثير نوع من الأقراص المنومة على الإنسان، أخذ طبيب مجموعتين من الأشخاص حجمهما على التوالي 50 و 100؛ أعطيت المجموعة الأولى أقراص النوع الجديد، فكان متوسط النوم هو 7.82 ساعة بانحراف معياري 0.24 ساعة؛ في حين أعطيت المجموعة الثانية الأقراص العادية (التقليدية) فكان متوسط النوم هو 6.75 ساعة بانحراف معياري 0.30 ساعة.

- أوجد مجالات الثقة 90 % و 99 % للفرق بين متوسط ساعة النوم الناتجة عن استخدام النوعين من الأقراص.

### التمرين الثالث:

نريد تقدير النسبة  $P$  لمجموعة من السلع الاستهلاكية التي تحمل علامة مميزة مأخوذة من مجتمع حجمه 50 مليون وحدة إنتاجية.

(أ) - نختبر أولا 100 وحدة استهلاكية، أخذت عشوائيا، ووجدنا أن 5 وحدات تحمل علامة مميزة.

- أوجد مجال الثقة للنسبة  $P$  ( $\alpha = 5\%$ ).

(ب) - نفرض أن المجال المحسوب في الفقرة أ معتبرا، من أجل ذلك نأخذ عينة معتبرة كذلك، ولنفرض أن النسبة  $P_0$  هي نفسها المحسوبة في الفقرة أ.



- ما هو حجم العينة اللازم أخذه إذا أردنا تقدير النسبة  $P$  ( $\alpha = 1\%$ )،
- أوجد عندئذ احتمال المجال الجديد  $I'$  حيث أن  $I' = 2I$ .

#### التمرين الرابع:

يرغب مدير مؤسسة في تقدير متوسط عدد الساعات التي يأخذها العمال لإنجاز عمل ما في حدود  $\pm 3$ ، بمجال ثقة 90%. فإذا كان الانحراف المعياري يساوي 15، فما هو حجم العينة اللازم أخذه (عدد الساعات).

#### التمرين الخامس:

ماكينة صناعية مختصة في إنتاج قطع حديد دائرية قطرها  $d$ . فإذا كان قطر القطع يخضع للتوزيع الطبيعي بانحراف معياري 1 مم. أخذت عينة عشوائية من منتج الماكينة حجمها 9 وتم قياس الأقطار المنتجة فكانت النتائج كما يلي (بالملم): 20.1، 19.9، 20، 19.8، 19.7، 20.2، 20.1، 23.1، 22.8. أوجد مجال الثقة 95% لمتوسط أقطار القطع المنتجة؟.

#### التمرين السادس:

أخذت عینتان حجمهما على التوالي 10 و 18 من مجتمعين موزعين طبيعيا. فإذا كان تباين العينة الأولى هو 30 و تباين العينة الثانية هو 25، فأوجد مجال الثقة 95% للنسبة بين تبايني المجتمعين اللذين أخذت منهما هاتان العينتان.



## الفصل الخامس

### نظرية الارتباط

#### مقدمة:

تنقسم الظواهر إلى قسمين: ظواهر مسببة وظواهر ناتجة. إن نظرية الارتباط، كفرع من فروع علم الإحصاء، تهتم بدراسة العلاقات الارتباطية المختلفة بين الظواهر المسببة والظواهر الناتجة، بهدف الوصول إلى صيغة رياضية (نموذج رياضي) تعبر عن العلاقة المفروضة. وتتجلى أهمية هذه النظرية في المجالات الاقتصادية المختلفة ومجالات العلوم الطبيعية، حيث تقدم الأدوات الرياضية اللازمة لإتخاذ القرارات المناسبة.

#### 1.5- أنواع العلاقات الارتباطية:

نرمز بـ  $X$  للظاهرة المسببة وبالرمز  $Y$  للظاهرة الناتجة. وتكون العلاقة بينهما إما على شكل:

##### • علاقات تابعة وارتباطية:

حيث يكون مقابل كل متغير مستقل  $X$  قيمة مقابلة لـ  $Y$  (المتغير المرتبط). فمثلا الصيغة  $L = 2\pi R$  هي علاقة ارتباطية بين محيط الدائرة ونصف قطرها  $R$ ، لأن  $\pi$  مقدار ثابت.

##### • علاقة طردية أو عكسية:

إن كلا من العلاقة الارتباطية بين ظاهرتين يمكن أن تكون طردية (كلما ازدادت قيم  $X$  ازدادت قيم  $Y$ )، أو عكسية (أي إذا ازدادت قيم  $X$  نقصت قيم  $Y$ ).

## • علاقات مستوية أو منحنية:

تكون العلاقة الارتباطية إما مستقيمة، تكون فيها تزايد الظاهرة  $Y$  على تزايد الظاهرة  $X$  ثابتة، أو منحنية حيث تكون هذه النسبة غير ثابتة.

### 2.5- الارتباط البسيط:

إن الارتباط البسيط يختص فقط في البحث عن العلاقة الارتباطية بين ظاهرتين فقط، أحدهما تكون الظاهرة الناتجة  $Y$  و الأخرى الظاهرة المسببة  $X$ . ينقسم الارتباط البسيط إلى نوعين أساسيين، هما الارتباط المستقيم و الارتباط المنحني.

### 1.2.5- الارتباط المستقيم و تمثيله:

يعبر عن الارتباط المستقيم بين المؤشرين  $X$  و  $Y$  بمعادلة مستقيم من الشكل:

$$Y^* = a_0 + a_1 X \quad (5.1)$$

وللتأكد من أن هذه العلاقة هي إرتباط مستقيم، فإنه يجب رسم شكل الانتشار للنقاط التي إحداثياتها  $(X_i, Y_i)$  الذي يجب أن يأخذ شكل الخط المستقيم.

### 1.1.2.5- طريقة المربعات الصغرى:

لتكن السلسلة الارتباطية للظاهرتين  $X$  و  $Y$  لكل منهما القيم الحقيقية التالية:

$$X = \{x_1, x_2, \dots, x_i\}$$

$$Y = \{y_1, y_2, \dots, y_i\}$$

إن القيم الممثلة بواسطة معادلة المستقيم

$$Y^* = a_0 + a_1 X$$

تسمى بالقيم التقديرية، و هي تلك الواقعة على خط المستقيم، بتعويض قيم  $X_i$  في المعادلة.

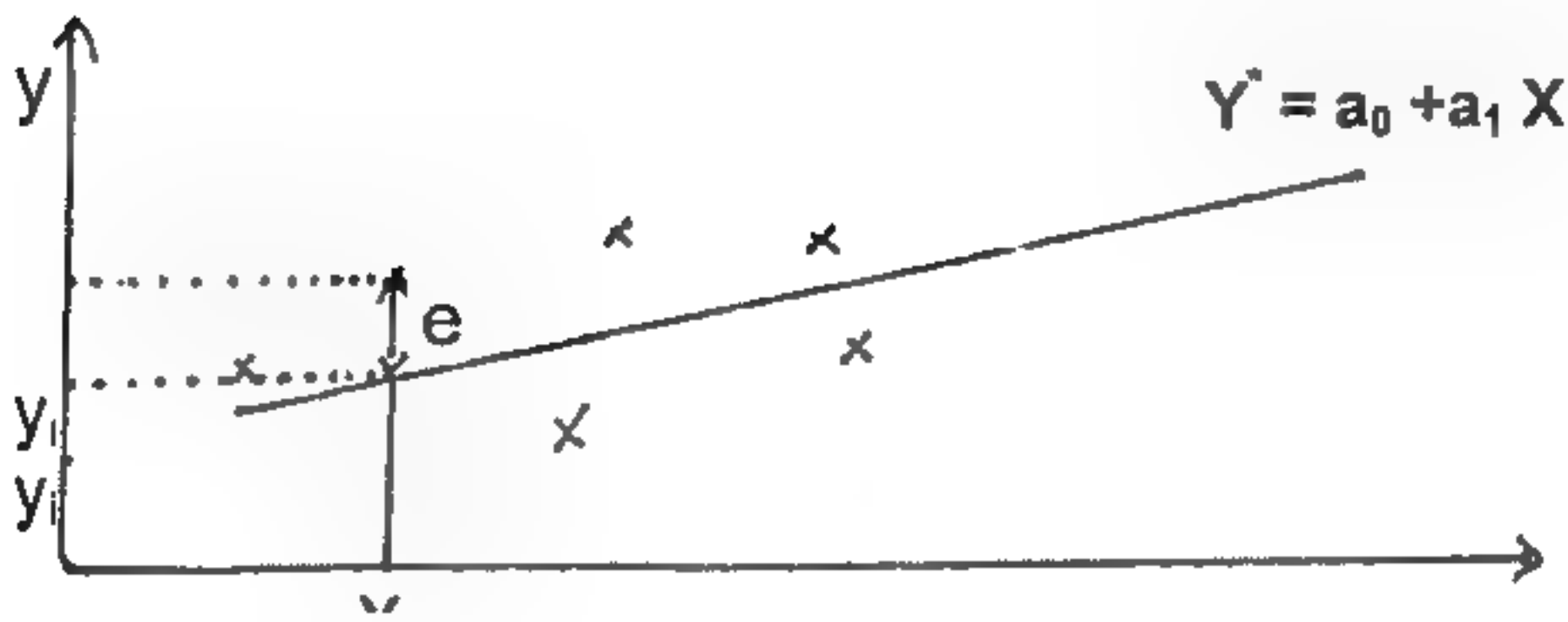
إن المعادلة (5.1) تعكس العلاقة الارتباطية بشكل تقريبي، لذلك فإنه من الطبيعي وجود سلسلة من المفروقات بين القيم الحقيقية والقيم التقديرية، وعليه يمكن تمثيل المعادلة (3.1) بالشكل التالي:

$$Y^* = a_0 + a_1 X + \varepsilon \quad (5.2)$$

$$\sum \varepsilon_i = 0$$

حيث:

من هذا نكتب:



الاتحاد المستقيم

$$Y^* = f(x) + \varepsilon$$

$$X_i = \{x_1, x_2, \dots, x_i\}$$

$$Y_i = \{y_1, y_2, \dots, y_i\}$$

$$\varepsilon_i = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i\}$$

إن المبدأ الأساسي لطريقة المربعات الصغرى يكمن في كون:

$$\varepsilon = \sum (y_i - y_i^*) = 0$$

أي:

$$\sum (y_i - y_i^*) \longrightarrow \text{Min} \quad (5.3)$$

والتي تعبر عن مجموع مربعات فروق القيم الحقيقية عن القيم التقديرية.  
لدينا:

$$\begin{aligned} \sum (y_i - y_i^*) &= 0 \\ \sum (y_i - a_1 x_i - a_0) &= 0 \\ \sum y_i - \sum a_1 x_i - \sum a_0 &= 0 \\ \sum y_i - a_1 \sum x_i - n a_0 &= 0 \end{aligned}$$

حجم العينة:  $n$

$$\begin{aligned} \sum y_i - a_1 \sum x_i = n a_0 &\Rightarrow a_0 = \frac{\sum y_i}{n} - a_1 \frac{\sum x_i}{n} \\ &\Rightarrow a_0 = \bar{y} - a_1 \bar{X} \dots \dots \dots (3.4) \end{aligned}$$

$$\begin{aligned} \sum (y_i - y_i^*)^2 &= \sum (y_i - a_1 x_i - a_0)^2 \\ &= \sum (y_i - a_1 x_i - \bar{y} + a_1 \bar{X})^2 \\ &= \sum [(y_i - \bar{y}) - a_1 (x_i - \bar{X})]^2 \end{aligned}$$

حيث أن:  $\bar{X}$  المتوسط الحسابي لقيم  $X$  و  $\bar{Y}$  يمثل المتوسط الحسابي لقيم  $Y$ .

بالتعويض نجد قيمة

$$\begin{aligned} a_1 &= \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \\ &= \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\frac{n}{N}} = \frac{\text{Cov}(x, y)}{\text{Var}(x)} \end{aligned} \tag{5.4}$$

حيث أن  $Cov(x,y)$  هو التباين المشترك أو التغاير و  $Var(x)$  يمثل تباين قيم  $X$ .

• طريقة أخرى لحساب  $a_0$  و  $a_1$ :  
لدينا:

$$Y_i^* = a_0 + a_1 X_i$$

نجمع وفق قيم  $i$  فتحصل على ما يلي:

$$\sum Y_i = na_0 + a_1 \sum X_i \dots\dots\dots 1$$

نضرب المعادلة الأصلية في  $X_i$  ثم نجمع وفق قيم  $i$ :

$$\sum X_i Y_i = a_0 \sum X_i + a_1 \sum X_i^2 \dots\dots 2$$

تشكل لدينا جملة معادلة ذات مجهولين  $a_0$  و  $a_1$ .

$$\left. \begin{array}{l} \sum Y_i = na_0 + a_1 \sum X_i \\ \sum X_i Y_i = a_0 \sum X_i + a_1 \sum X_i^2 \end{array} \right\} \quad (5.5)$$

تحل هذه الجملة بإحدى الطرق الرياضية (المقارنة، التعويض أو المصفوفات).

ملاحظة: إذا قسمنا أطراف المعادلة (5.5) على  $n$  فإننا نحصل على ما يلي:

$$\bar{Y} = a_0 + a_1 \bar{X} \quad (5.6)$$

أي أن المستقيم الذي نبحث عنه يمر من النقطة  $(\bar{X}, \bar{Y})$ . هذه الخاصية يتميز بها الارتباط المستقيم.

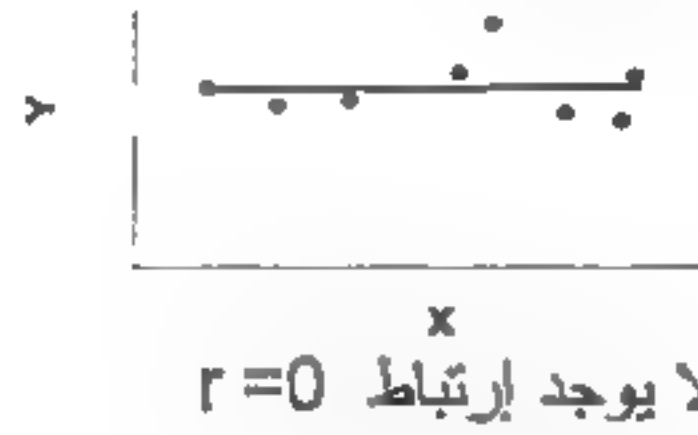
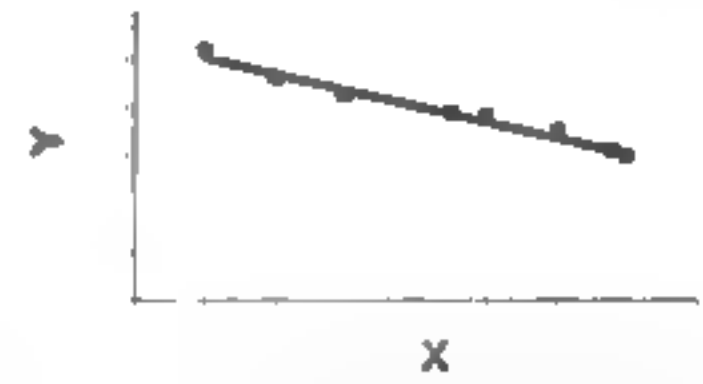
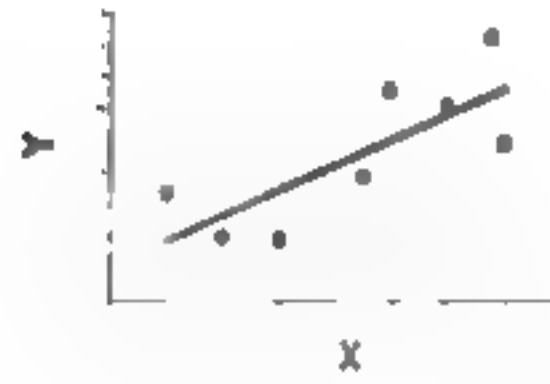
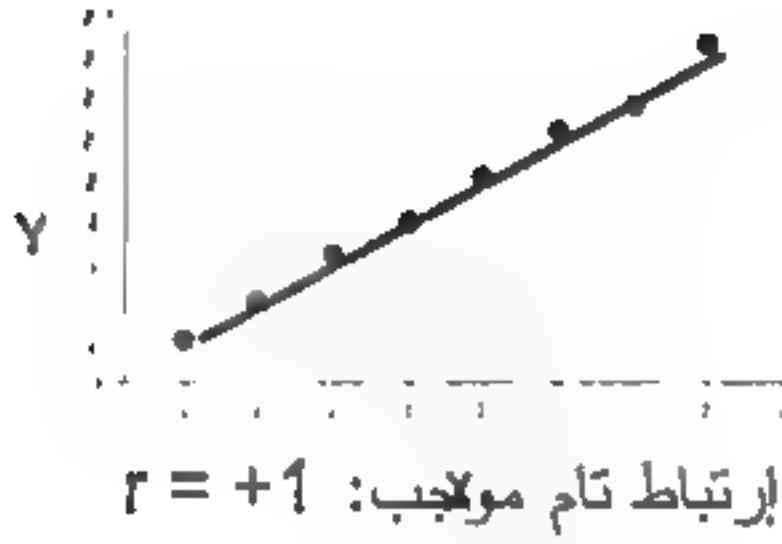
### 2.1.2.5- معامل الارتباط:

معامل الارتباط هو معامل يقيس شدة العلاقة الارتباطية بين المتغيرين X و Y. تتراوح قيمته ما بين -1 و +1، و يعطى بالعلاقة التالية:

$$r = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(x)\text{Var}(y)}} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}} = \frac{\text{Cov}(x, y)}{n\sigma_x \cdot \sigma_y} \quad (5.7)$$

حيث:  $\sigma_x$  و  $\sigma_y$  هما الانحرافات المعيارية لـ X و Y على التوالي.  
إذا كان:

- $r = +1$  هذا يعني وجود ارتباط تام موجب (علاقة طردية)،
- $r = -1$  يعني ارتباط تام سالب (علاقة عكسية)،
- $r = 0$  لا يوجد ارتباط (استقلال).



• عبارة أخرى لمعامل الارتباط:

بغية تسهيل الحسابات يمكن تبسيط عبارة معامل الارتباط لتأخذ الصيغة التالية:

$$r = \frac{\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (5.8)$$



### ➤ مثال:

لدى دراسة 8 مؤسسات لصناعة النسيج إستخلصنا عدد الوحدات الإنتاجية في كل مؤسسة X، و متوسط الإنتاج اليومي فيها Y (بالطن).  
تحصلنا على النتائج التالية:

X	10	20	30	40	50	60	70	80
Y	21	40	62	78	100	122	135	164

المطلوب دراسة العلاقة الارتباطية بين عدد الوحدات الإنتاجية و متوسط الإنتاج اليومي.

### ➤ الحل:

من المقارنة الأولية، يظهر وجود علاقة طردية بين المتغيرين، و يظهر ذلك جليا بعد رسم شكل الانتشار، الذي يأخذ إتجاه إرتباط مستقيم، و عليه فإن المعادلة المقترحة هي كما يلي:

$$Y_i^* = a_0 + a_1 X_i$$

ولحساب قيم ثوابت المعادلة، نملأ الجدول التالي:

X	Y	XY	X <sup>2</sup>	Y <sup>2</sup>
10	21	210	100	441
20	40	800	400	1600
30	62	1860	900	3844
40	78	3120	1600	6084
50	100	5000	2500	10000
60	122	7320	3600	14884
70	135	9450	4900	18225
80	164	13120	6400	26896
360	722	40880	20400	81974

إنطلاقا من المعادلة (5.5) نجد:

$$\begin{aligned} 8 a_0 + 360 a_1 &= 722 \\ 360 a_0 + 20400 a_1 &= 40880 \end{aligned}$$

بعد حل هذه الجملة بإحدى الطرق الرياضية، تكون المعادلة كما يلي:

$$Y_i^* = 0.36 + 2 X_i$$

ويمكن بسهولة حساب معامل الارتباط  $r$  من خلال العبارة (5. 7) فنجد:

$$r = \frac{8(40880) - (360)(722)}{\sqrt{[8(20400) - (360)^2][8(81974) - (722)^2]}} = 0.998$$

### 3.1.2.5- اختبار دلالة معامل الارتباط:

ليكن  $\rho$  معامل ارتباط مجتمع ما، وليكن  $r$  معامل الارتباط لعينة أخذت من هذا المجتمع. نريد أن نختبر إذا ما كان معامل ارتباط المجتمع يختلف معنويا عن الصفر (عند مستوى دلالة  $\alpha$ )، أي:

$$\begin{aligned} H_0: \rho &= 0 \\ H_1: \rho &\neq 0 \end{aligned}$$

لإجراء هذا الاختبار فإننا نستخدم توزيع  $t$ ، فتكون قيمة  $t$  كما يلي:

$$t = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2} \quad (5.9) \quad \text{عند درجات حرية } n-2$$

فإذا كانت  $t$  التطبيقية  $<$  من  $t$  النظرية، فإننا نرفض  $H_0$  لصالح  $H_1$ .

➤ مثال:

أخذنا عينة حجمها 18 مفردة، وقمنا بدراسة علاقة ارتباطية بين ظاهرتين في هذه العينة، فوجدنا أن معامل الارتباط يساوي 0.32. هل يمكن أن يكون معامل الارتباط في مجتمع العينة أكبر معنويا من الصفر.

➤ الحل:

$$t_{\text{التطبيقية}} = \frac{0.32 \sqrt{18-2}}{\sqrt{1-(0.32)^2}} = 1.35$$

من جدول توزيع  $t$  (عند 0.95) يمكن بسهولة قراءة قيمة  $t$  النظرية، فنجد القيمة 1.75. و حيث أن  $t$  النظرية أكبر من  $t$  التطبيقية، فإننا نقبل  $H_0$ ، أي أن معامل ارتباط المجتمع لا يختلف معنويا عن الصفر.

• ملاحظة هامة:

من أجل الحصول على علاقة إرتباطية لها معنوية عند  $\alpha$ ، يجب رفع حجم العينة، في هذه الحالة فإن العلاقة الإرتباطية تتبع التوزيع الطبيعي، وبذلك يكون معامل إرتباط المجتمع  $\rho$  يختلف معنويا عن الصفر.

لإجراء هذا الاختبار، فإننا نعرف الدالة الإحتمالية التالية:

$$X = \frac{1}{2} \ln \frac{(1+r)}{(1-r)} = 1.1513 \log \frac{1+r}{1-r} \quad (5.10)$$

إن الدالة الاحتمالية  $X^*$  تخضع للتوزيع الطبيعي الذي وسطه الحسابي  $\mu_x$  وانحرافه المعياري  $\sigma_x$  حيث:

$$\mu_x = \frac{1}{2} \ln \frac{(1+\rho)}{(1-\rho)} \quad , \quad \sigma_x = \frac{1}{\sqrt{n-3}}$$

➤ مثال:

عينة حجمها 24 مفردة، معامل إرتباطها 0.75. هل يمكن أن نرفض الفرضية القائلة بأن معامل الإرتباط للمجتمع الذي أخذت منه العينة أقل من 0.6، عند مستوى دلالة 5%.

➤ الحل:

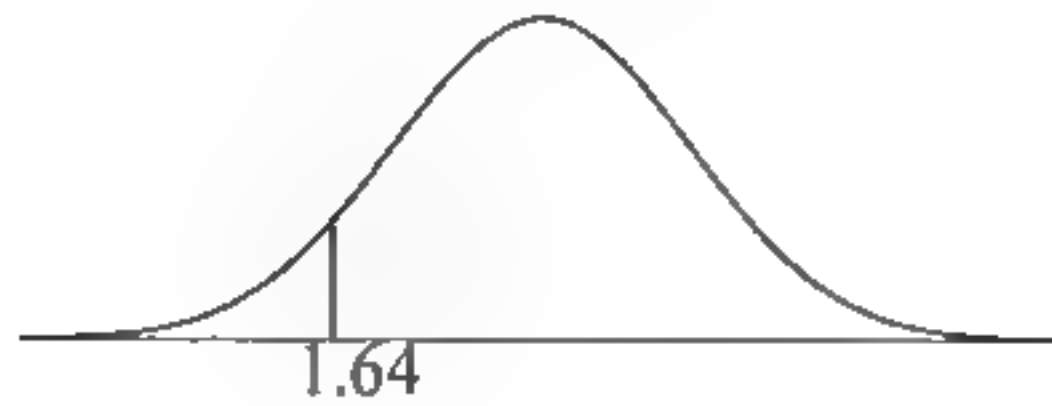
لدينا:

$$H_0: \rho = 0.60$$

$$H_1: \rho < 0.60$$

$$X = \frac{1}{2} \ln \frac{(1+r)}{(1-r)} = \frac{1}{2} \ln \frac{(1+0.75)}{(1-0.75)} = 0.9730$$

وحيث أن الدالة الاحتمالية  $X$  تخضع للتوزيع الطبيعي الذي وسطه الحسابي  $\mu_x$



والانحراف المعياري  $\sigma_x$ ، فإن:

$$\mu_x = \frac{1}{2} \ln \left( \frac{1 + 0.6}{1 - 0.6} \right) = 0.69$$

$$\sigma_x = \frac{1}{\sqrt{24 - 3}} = 0.21$$

نقوم بتحويل الوحدة  $X$  إلى الوحدة  $Z$  وفق العلاقة:

$$Z = \frac{X - \mu}{\sigma_x} = \frac{0.9730 - 0.69}{0.21} = 1.34$$

وحيث أن  $Z = 1.64$  النظرية أكبر من  $Z = 1.34$  التطبيقية، فإننا نقبل الفرضية  $H_0$ ، أي أن  $p = 0.60$

● حالة إختبار تساوي معاملي إرتباط:

لتكن العينتان  $A_1$  و  $A_2$  حجمهما على التوالي  $n_1$  و  $n_2$ ، معاملي إرتباطهما هما على التوالي  $r_1$  و  $r_2$ . نريد أن نختبر الفرضية:

$$H_0: r_1 = r_2$$

لإجراء هذا الإختبار، يتحتم علينا حساب  $Z_1$  و  $Z_2$  المقابلتين لـ  $r_1$  و  $r_2$  باستخدام المعادلة المعطاة في (5.10) لحساب قيمة  $X$  بدلالة معامل الإرتباط

المفترض. إن حساب قيمتي  $Z_1$  و  $Z_2$  من خلال حساب معامل ارتباط محول  $\hat{r}$  وفق العلاقة التالية:

$$\hat{r} = \frac{e^{2x} - 1}{e^{2x} + 1}, \text{ حيث أن } \hat{Z} \text{ المقابلة هي } \hat{Z} = \frac{1}{2} \ln \frac{(1+\hat{r})}{(1-\hat{r})}$$

يسمح الجدول رقم 8 (أنظر الملحق) بإعطاء قيم  $Z$  دون اللجوء إلى هذه الحسابات.

نحسب إحصائية هذا الاختبار وفق العلاقة التالية:

$$Z = \frac{Z_1 - Z_2}{\sigma_{Z_1 - Z_2}} \quad (5.11)$$

حيث:

$$\sigma_{Z_1 - Z_2} = \sqrt{\sigma^2_{Z_1} + \sigma^2_{Z_2}} = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}} \quad (5.12)$$

نرفض الفرضية العدمية  $H_0$  عند مستوى دلالة  $\alpha$  عندما يكون:

$$Z_{\text{obs}} \geq 1,96 \quad (\alpha = 5\%)$$

➤ مثال:

قمنا بدراسة العلاقة الارتباطية بين نسبة الكربون C و نسبة الأزوت N في التربة، من خلال عينتين، حجمهما  $n_1=10, n_2=11$  مأخوذتين من منطقتين مختلفتين. وجدنا معاملات الارتباط على التوالي:  $r_1=0,349$  ;  $r_2=0,827$  - هل توجد فروق معنوية لمعاملات الارتباط بين العينتين؟.

➤ الحل:

$$H_0: r_1 = r_2$$

[من الجدول 8، أنظر الملحق]  $Z_1 = 0,3643$  ,  $Z_2 = 1,1784$

$$Z_{obs} = \frac{|0,364 - 1,178|}{\sqrt{\frac{1}{11} + \frac{1}{12}}} = 1.952$$

وحيث أن:  $Z_{obs} \leq 1,96$  فإننا نقبل  $H_0$  عند مستوى دلالة  $\alpha = 5\%$ .

• اختبار تساوي أكثر من معاملي ارتباط:

يمكن تعميم هذا الاختبار ليأخذ  $p$  معاملات ارتباط، حيث يكون:

$$(H_0: r_1 = r_2 = \dots = r_p)$$

إن هذا الاختبار يخضع لتوزيع  $\chi^2$ :

$$\chi_{obs}^2 = \sum_{i=1}^p (n_i - 3)(z_i - \bar{z})^2 \quad (5.13)$$

لقيم  $z_i$

نرفض الفرضية العدمية إذا كانت قيمة الـ  $\chi_{obs}^2$  التطبيقية أكبر من  $\chi_{ddl\alpha}^2$  النظرية

$$= p - 1 \text{ درجات الحرية}$$

وعند تساوي أحجام العينات  $n_1 = n_2 = \dots = n_p = n$  فإن:

$$\chi_{obs}^2 = (n - 3) \sum_{i=1}^p (z_i - \bar{z})^2 \quad (5.14)$$

ويمكن حساب المتوسط  $\bar{z}$  بسهولة، حيث:

$$\frac{Z_1 + Z_2 + \dots + Z_n}{n}$$

### ➤ مثال:

قمنا بدراسة العلاقة الارتباطية بين التداول الاقتصادي لسلعة ما وبين قيمة هذه السلعة في أربع مناطق مختلفة، من خلال أخذ 10 عينات، فكانت النتائج كما يلي:

$$r_1 = 0,349 \quad , \quad r_2 = 0,827 \quad , \quad r_3 = 0,667 \quad , \quad r_4 = 0,807$$

- هل توجد فروق معنوية لمعاملات الارتباط؟.

### ➤ الحل:

$$H_0: r_1 = r_2 = r_3 = r_4$$

من الجدول (أنظر الملحق) نحسب القيم التالية:

$$Z_1 = 0,3643 \quad , \quad Z_2 = 1,1784 \quad , \quad Z_3 = 0,8054 \quad , \quad Z_4 = 1,1183$$

فيكون المتوسط:  $\bar{Z} = 0,8666$

$$\chi_{obs}^2 = (n - 3) \sum (Z_i - 0.86)^2 = (10 - 3) \cdot 0,4166 = 2,92$$

$$dII = 4 - 1 = 3 \quad \text{و} \quad \alpha = 5\%$$

$$\chi^2 = 7,81$$

وحيث أن  $\chi_{obs}^2 > \chi_{0,05,3}^2$  فإننا نقبل  $H_0$ . عند مستوى معنوية 5%، هذا يعني تساوي معاملات الارتباط في المناطق الأربع.

### 4.1.2.5- معامل التحديد $R^2$ :

يقيس معامل التحديد كذلك شدة العلاقة الارتباطية على نفس الطريقة التي يعمل بها معامل الارتباط  $r$ ، وهو يعرف بأنه حاصل قسمة تباين القيم التقديرية على تباين القيم الحقيقية، أي أن:

$$R^2 = \frac{\sigma_y^{*2}}{\sigma_y^2} \quad , \quad \text{حيث أن } 0 < R^2 < 1 \quad (5.15)$$

### 5.1.2.5- دراسة الخطأ المرتكب و تقديره:

عندما نمثل العلاقة الارتباطية بين المتغيرين  $X$  و  $Y$  بواسطة معادلة رياضية ما فإننا نحصل على القيم التقديرية  $Y_i^*$ ، و إن هذه القيم لا تنطبق على القيم الحقيقية. يسمى الفرق بين القيم الحقيقية و القيم النظرية بقيم البواقي، و نرمز له بالرمز  $\varepsilon_i$ ، حيث:

$$\varepsilon_i = Y_i - Y_i^* \quad (5.16)$$

إن المقدار  $\varepsilon_i$  (عدد حقيقي) ناجم عن تأثير عدة أخطاء متراكمة أهمها:

- خطأ القياس: أثناء قياس قيم  $X$  و قيم  $Y$ ،
- خطأ العينة: عدم شمولية المعلومات و الاعتماد على قيم العينات،
- الخطأ في اختيار معادلة التمثيل المناسبة،
- الأخطاء الغير المفسرة الناتجة عن وجود عوامل أخرى لم تكن في الحسبان وتدخل في معادلة التمثيل،
- الأخطاء العشوائية، و هي الأخطاء التي تظهر فجأة في مرحلة ما من مراحل تطور العلاقة الارتباطية.

ولذلك فإننا نفترض أن  $\varepsilon_i$  هو متغير عشوائي يتبع التوزيع الطبيعي  $N(0, S_{yy}^2)$  حيث أن:

$$S_{yy}^2 = \sigma_{\varepsilon}^2 = \frac{1}{n} \sum (\varepsilon_i - 0)^2 = \frac{1}{n} \sum (Y_i - Y_i^*)^2 \quad (5.17)$$

واعتمادا على جدول التوزيع الطبيعي فإن:

$$P(-1 \leq Z \leq +1) = 0.68$$

$$P(-1 \leq \frac{\varepsilon_i}{\sigma_{\varepsilon}} \leq +1) = 0.68$$

$$P(-\sigma_{\varepsilon} \leq \varepsilon_i \leq +\sigma_{\varepsilon}) = 0.68$$

$$P(-\sigma_{\varepsilon} \leq Y_i - Y_i^* \leq +\sigma_{\varepsilon}) = 0.68$$



وحيث أن:  $\sigma_{\varepsilon i}^2 = S_{yy}^*$  فإننا نجد:

$$P(-S_{yy}^* \leq Y_i - Y_i^* \leq S_{yy}^*) = 0.68$$

$$P(Y_i^* - S_{yy}^* \leq Y_i \leq Y_i^* + S_{yy}^*) = 0.68 \quad \text{أي أن:}$$

مما يعني أن القيم الحقيقية محصورة في المجال:

$$[Y_i^* - S_{yy}^*, Y_i^* + S_{yy}^*] \quad \text{بإحتمال قدره } 0.68 \quad (5.18)$$

و بالمثل نجد:

$$[Y_i^* - 2S_{yy}^*, Y_i^* + 2S_{yy}^*] \quad \text{بإحتمال قدره } 0.96 \quad (5.19)$$

$$[Y_i^* - 3S_{yy}^*, Y_i^* + 3S_{yy}^*] \quad \text{بإحتمال قدره } 0.997 \quad (5.20)$$

تسمى هذه المجالات بمجالات الثقة للقيم الحقيقية بإحتمال ثقة  $(1 - \alpha)$

• عبارة أخرى لمعامل التحديد  $R^2$ :

مما سبق نستنتج أن  $\sigma_{Yi}^2 = \sigma_{Y^*i}^2 + S_{yy}^*$  : (التباين الكلي = تباين القيم التقديرية - تباين الخطأ)،

أي أن:  $\sigma_{Y^*i}^2 = \sigma_{Yi}^2 - S_{yy}^*$  ، من جهة أخرى لدينا:

$$R^2 = \frac{\sigma_{Y^*}^2}{\sigma_Y^2} = 1 - \frac{S_{yy}^*}{\sigma_Y^2} \quad (5.21)$$

- إذا كان  $R^2 = 1$  فإن المقدار  $\frac{S_{yy}^*}{\sigma_Y^2} = 0$  ، أي أن  $S_{YY}^* = 0$  ، وبالتالي

تنطبق القيم التقديرية على القيم الحقيقية و هو ما يعني إرتباط تام،

- إذا كان  $R^2 = 0$  فإن المقدار  $\frac{S_{yy}^*}{\sigma_Y^2} = 1$  ، أي أن  $S_{YY}^* = \sigma_Y^2$  ،

وبالتالي تباين قيم الخطأ هو نفسه تباين القيم الحقيقية وهو ما يعني إستقلال تام.

• العلاقة بين معامل الارتباط  $r$  و معامل التحديد  $R^2$ :

تتحقق هذه العلاقة فقط عندما يكون الارتباط خطي، فنحصل على:

$$R^2 = r^2$$

أما في باقي الحالات الارتباطية فإنه لا يمكن إستنتاج هذه العلاقة.

6.1.2.5- إختبار معامل التحديد  $R^2$ :

يخضع هذا الإختبار إلى توزيع  $F$  (فيشر)، و الذي يحسب من خلال

العلاقة التالية:

$$F = \frac{\frac{\sigma^2 y^*}{S^2_{yy}^*} \cdot \frac{n - m}{m - 1}}{\text{التطبيقية}} \quad (5.22)$$

حيث أن  $m$  هي عدد الثوابت المقدرة في معادلة التمثيل،  $n$  هو حجم العينة في السلسلة الارتباطية.

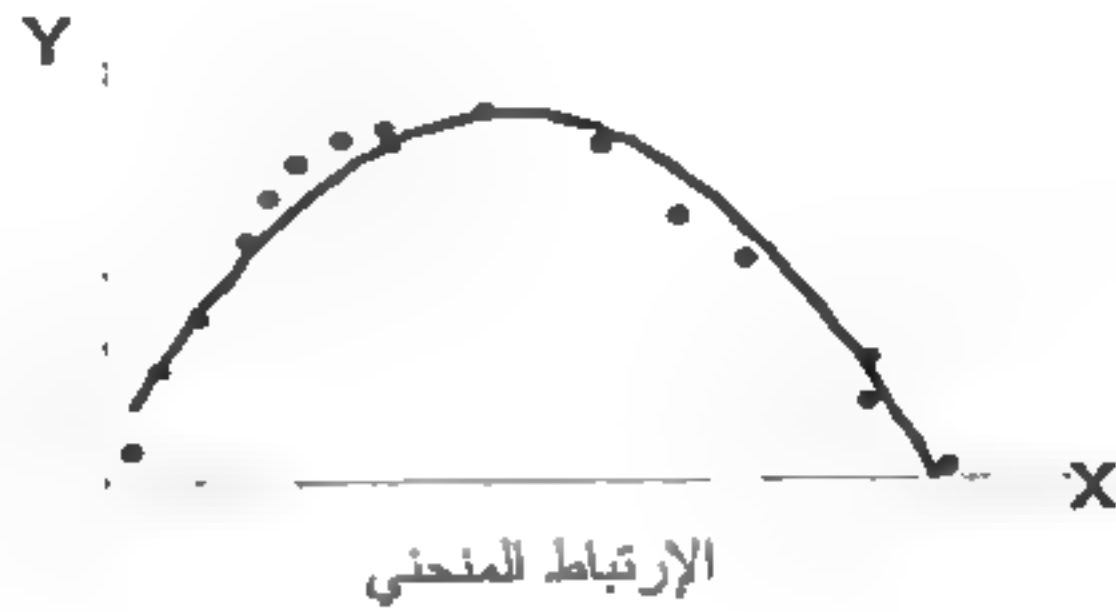
في جدول  $F$  يمكن أن نقرأ  $V_1 = m-1$  و  $V_2 = n-m$  و من ثم قيمة  $F$  النظرية. فإذا كانت  $F$  النظرية أكبر من  $F$  التطبيقية فإن قيمة  $R^2$  ليس لديها معنوية (عند مستوى معنوية  $\alpha$ )، لذلك نضطر إلى تغيير معادلة التمثيل أو زيادة حجم العينة (أو كلاهما معا) للحصول مستوى معنوية لمعامل التحديد.

3.5- الارتباط المنحني:

إذا كان شكل الارتباط لا يأخذ شكل مستقيم، فإنه من الضروري البحث عن الشكل المنحني وعن معادلة التمثيل الخاصة به.

### 1.3.5- التمثيل بواسطة معادلة من الدرجة الثانية (معادلة القطع المكافئ):

إن هذا الارتباط يأخذ المعادلة التالية: (5.23)  $Y^* = a_0 + a_1X + a_2X^2$



ولحساب الثوابت نستخدم طريقة المربعات الصغرى فنجد:

$$a_0 = \frac{\sum (x_i - \bar{x})^2 \sum (x_i - \bar{x})^2 (y_i - \bar{y})}{[\sum (x_i - \bar{x})^2]^2 - N \sum (x_i - \bar{x})^4}$$

$$a_1 = \frac{\sum (x_i - \bar{x})^2 (y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$a_2 = \frac{-N.C}{\sum (x_i - \bar{x})^2}$$

يمكن استعمال كذلك طريقة المجاميع فنجد:

$$Y^* = a_0 + a_1X + a_2X^2$$

نجمع وفق قيم  $i$  فتحصل على ما يلي:

$$\sum Y_i = na_0 + a_1\sum X_i + a_2\sum X_i^2 \dots\dots\dots 1$$

نضرب المعادلة الأصلية في  $x_i$  ثم نجمع وفق قيم  $i$ :

$$\sum X_i Y_i = a_0\sum X_i + a_1\sum X_i^2 + a_2\sum X_i^3 \dots\dots 2$$

نضرب المعادلة الأصلية في  $X_i^2$  ثم نجمع وفق قيم  $i$ :

$$\sum X_i^2 Y_i = a_0\sum X_i^2 + a_1\sum X_i^3 + a_2\sum X_i^4 \dots\dots 2$$

تشكل لدينا جملة معادلة ذات 3 مجاهيل  $a_0$ ،  $a_1$  و  $a_2$ .

$$\left. \begin{aligned} \sum Y_i &= na_0 + a_1 \sum X_i + a_2 \sum X_i^2 \\ \sum X_i Y_i &= a_0 \sum X_i + a_1 \sum X_i^2 + a_2 \sum X_i^3 \\ \sum X_i^2 Y_i &= a_0 \sum X_i^2 + a_1 \sum X_i^3 + a_2 \sum X_i^4 \end{aligned} \right\} (5.24)$$

تحل هذه الجملة بإحدى طرق حساب المصفوفات.

### 2.3.5- التمثيل بواسطة معادلة من الدرجة الثالثة:

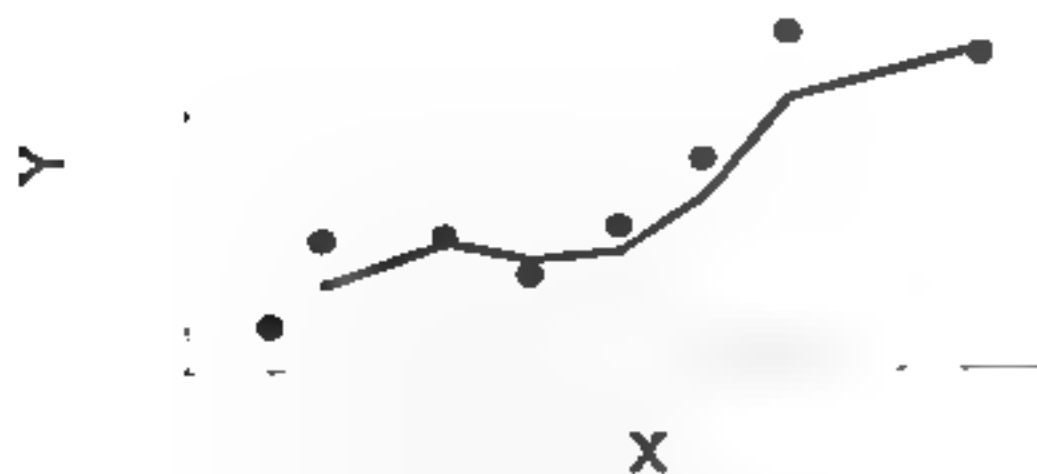
إن هذه المعادلة تأخذ الشكل التالي:

$$Y^* = a_0 + a_1 X + a_2 X^2 + a_3 X^3 \quad (5.25)$$

يمتاز منحنى هذا التمثيل بوجود نقطة إنعطاف (وجود تعرجات ذات قيمتين صغرى وعظمى). لحساب ثوابت هذه المعادلة نستخدم الأسلوب المستخدم في المعادلات السابقة فنجد:

$$\left. \begin{aligned} \sum Y_i &= na_0 + a_1 \sum X_i + a_2 \sum X_i^2 + a_3 \sum X_i^3 \\ \sum X_i Y_i &= a_0 \sum X_i + a_1 \sum X_i^2 + a_2 \sum X_i^3 + a_3 \sum X_i^4 \\ \sum X_i^2 Y_i &= a_0 \sum X_i^2 + a_1 \sum X_i^3 + a_2 \sum X_i^4 + a_3 \sum X_i^5 \\ \sum X_i^3 Y_i &= a_0 \sum X_i^3 + a_1 \sum X_i^4 + a_2 \sum X_i^5 + a_3 \sum X_i^6 \end{aligned} \right\} (5.26)$$

نتحصل في الأخير على جملة معادلة ذات أربعة مجاهيل، تحل من خلال حساب المصفوفات.

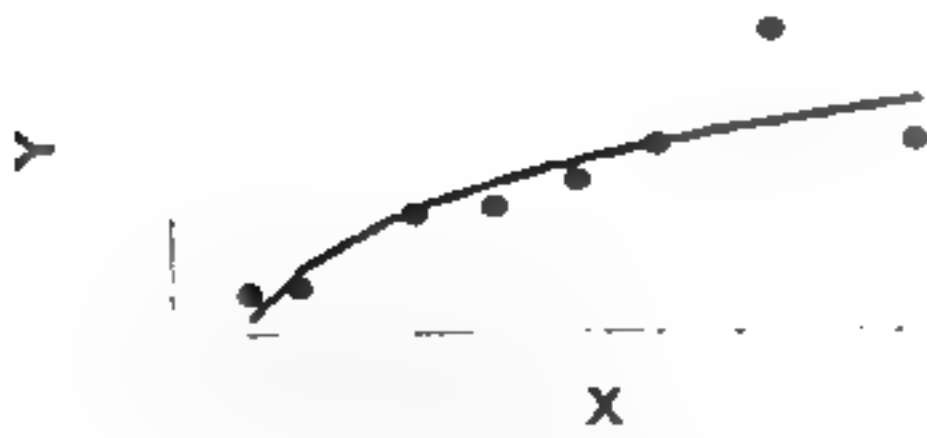


التمثيل بواسطة معادلة من الدرجة الثالثة

### 3.3.5- التمثيل بواسطة معادلة لوغاريتمية:

إن هذا الارتباط يأخذ المعادلة التالية:

$$Y^* = a_0 + a_1 \ln X, X > 0 \quad (5.27)$$



التمثيل بواسطة معادلة لوغاريتمية

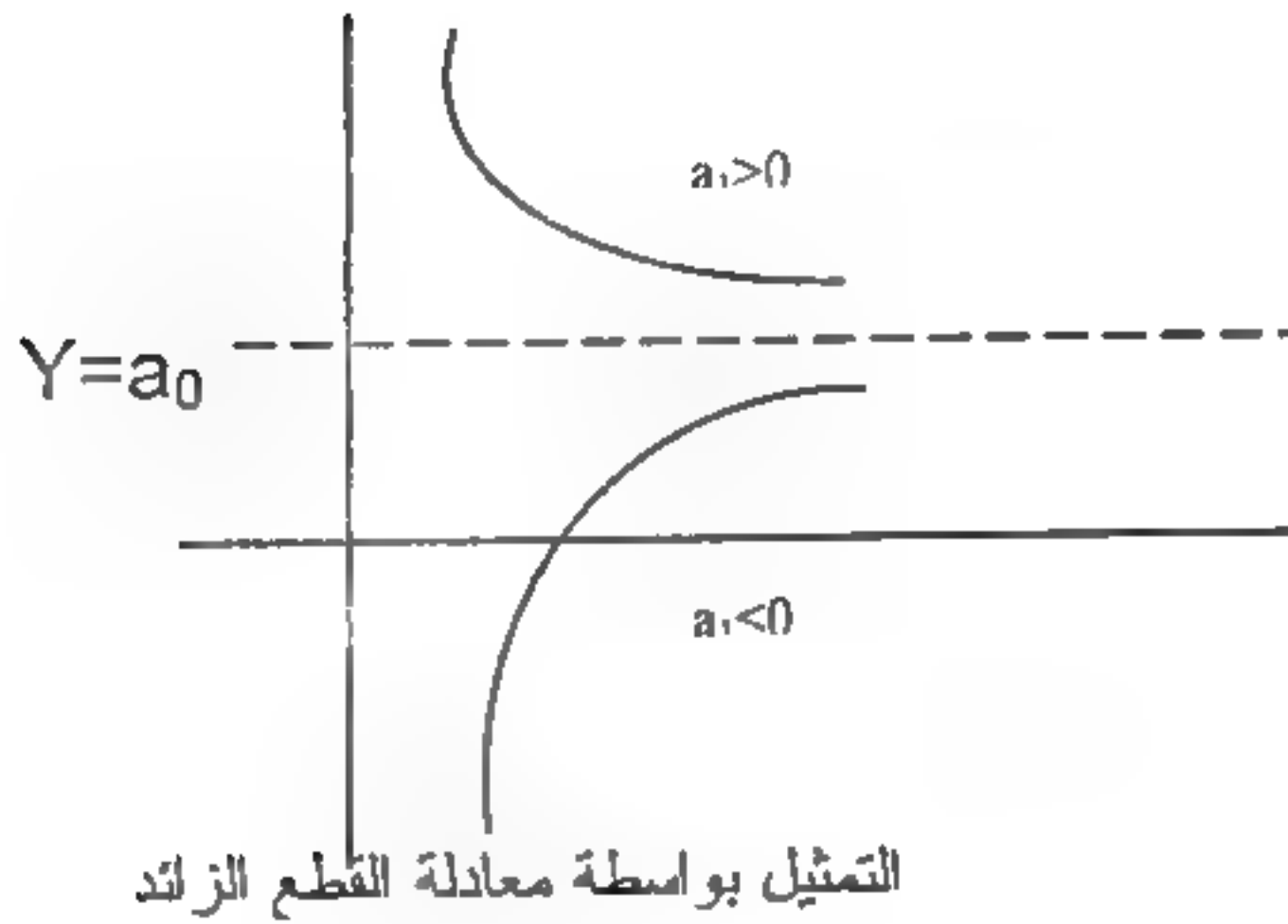
نلاحظ أنه كلما إزدادت قيم  $X$  تزداد قيم  $Y$  ولكن بتباطأ. لإيجاد الثابتين  $a_0$  و  $a_1$ ، نستخدم طريقة المربعات الصغرى فنحصل على:

$$\left. \begin{aligned} \sum Y_i &= na_0 + a_1 \sum \ln X_i \\ \sum Y_i \ln X_i &= a_0 \sum \ln X_i + a_1 \sum (\ln X_i)^2 \end{aligned} \right\} \quad (5.28)$$

وذلك بعد ضرب المعادلة الأصلية في  $\ln X$  ثم الجمع وفق قيم  $i$ .

### 4.3.5- التمثيل بواسطة معادلة القطع الزائد:

ويمكن أن نميز نوعين:



التمثيل بواسطة معادلة القطع الزائد

$$Y^* = a_0 + \frac{a_1}{X} \quad (5.29) \quad \text{— المعادلة من الشكل}$$

إن التمثيل البياني لهذه المعادلة يأخذ شكلين حسب إشارة  $a_1$ .  
من مميزات هذه المعادلة أنه إذا كان  $X$  ينتهي إلى المالا نهاية فإن  $Y$  تنتهي إلى الصفر. ولحساب قيم الثابتين  $a_0$  و  $a_1$  فإننا نستخدم طريقة المربعات الصغرى فنحصل على ما يلي:

$$\left. \begin{aligned} \sum Y_i &= na_0 + a_1 \sum \frac{1}{X} \\ \sum \frac{Y}{X} &= a_0 \sum \frac{1}{X} + a_1 \sum \frac{1}{X^2} \end{aligned} \right\} \quad (5.30)$$

وذلك بعد ضرب المعادلة الأصلية في  $\frac{1}{X}$  ثم الجمع وفق قيم  $i$ .

$$\text{— المعادلة من الشكل: } \frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1$$

يمكن كتابة هذه المعادلة على الشكل  $Y^2 = \frac{b^2}{a^2} X^2 - b^2$  ، ثم نفرض مايلي:

$$\beta = b^2 \text{ و } \alpha = \frac{b^2}{a^2} ، \text{ نتحصل على ما يلي:}$$

$$Y^2 = \alpha X^2 - \beta$$

وباستعمال طريقة المربعات الصغرى، نتحصل على الشكل التالي:

$$\left. \begin{aligned} \sum y^2 &= \alpha \sum x^2 - n \beta \\ \sum y^2 x^2 &= \alpha \sum x^4 - \beta \sum x^2 \end{aligned} \right\} \quad (5.31)$$

بعد الجمع وضرب المعادلة الأصلية في  $X^2$  ثم الجمع وفق قيم  $i$ .

بعد إيجاد قيمتي الثابتين  $\alpha$  و  $\beta$  و تعويضهما في المعادلة (3.31)، نقوم بحل هذه المعادلة فنحصل على جذرين:

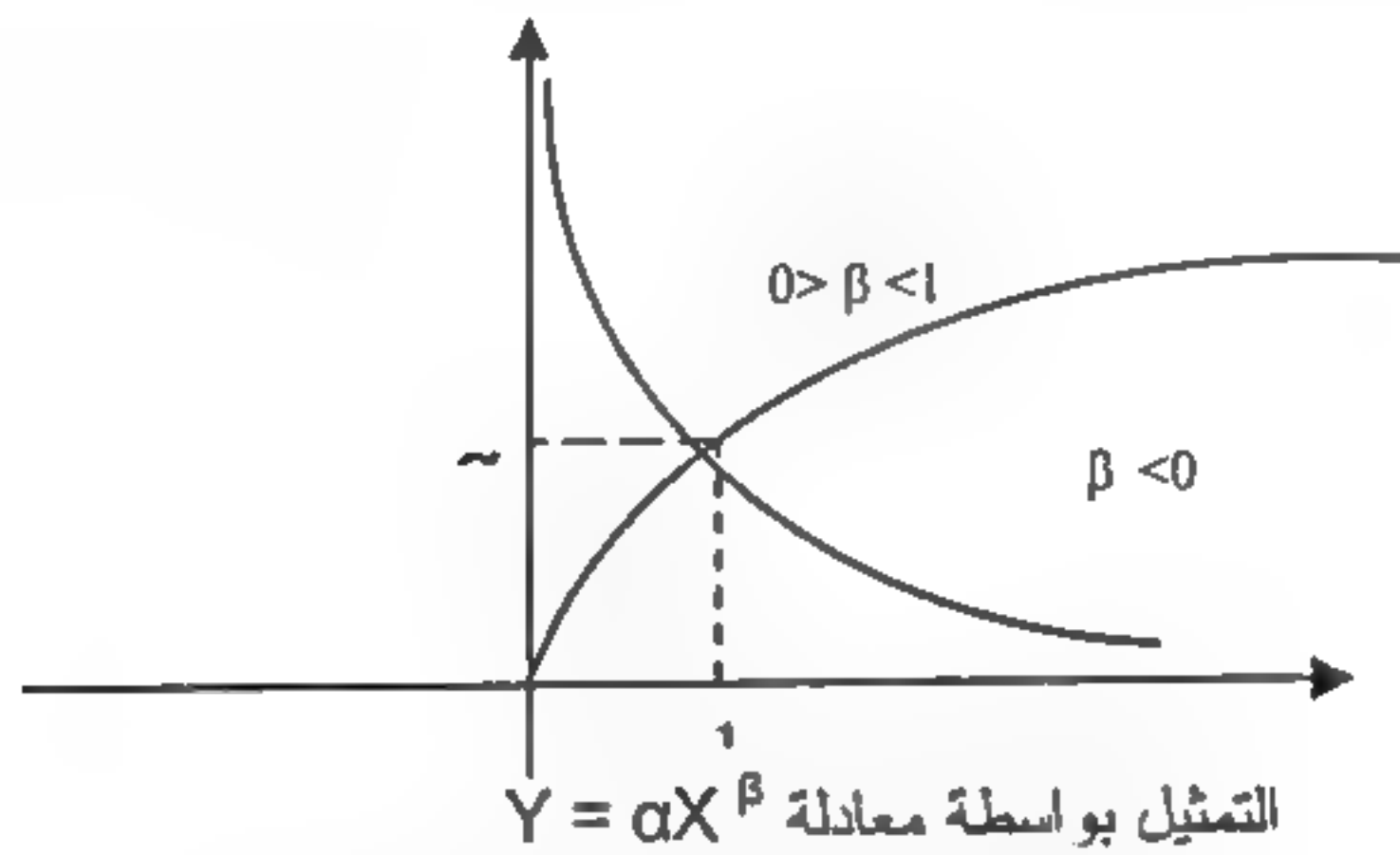
$$Y^* = +\sqrt{\alpha X^2 - \beta}$$

$$Y^* = -\sqrt{\alpha X^2 - \beta} \quad \text{حيث أن المقدار } \alpha X^2 - \beta \geq 0$$

يتم اختيار المعادلة المناسبة حسب المعطيات المقدمة (المعادلة الموجبة في المسائل الإقتصادية).

• التمثيل بواسطة المعادلة من الشكل:  $Y = \alpha X^\beta$ ,  $X > 0$

حيث أن  $\alpha$  و  $\beta$  عدنان حقيقيان. إن هذه المعادلة تأخذ شكلين حسب إشارة  $\beta$  كما هو موضح في الشكل التالي:



ولمعالجة هذه المعادلة، نأخذ لوغاريتم الطرفين:

$$\ln Y = \ln \alpha + \beta \ln X$$

بوضع:

$$\ln Y = Z$$

$$\ln \alpha = A$$

$$\ln X = X$$

تصبح المعادلة كما يلي:

$$Z = A + \beta X \quad (5.32)$$

نلاحظ أن المعادلة المتشكلة ما هي إلا معادلة خط مستقيم، و لذلك فإن إيجاد قيمتي الثابتين يخضع لنفس الطريقة المتبعة في الارتباط المستقيم. و بالرجوع إلى قيم اللوغارتم، حيث أن  $\ln Y = Z$  نتحصل على الجملة التالية:

$$\sum \ln Y_i = nA + \beta \sum X_i$$

$$\sum X_i \ln Y_i = A \sum X_i + \beta \sum X_i^2$$

أي أن:

$$\left. \begin{aligned} \sum \ln Y_i &= nA + \beta \sum \ln X \\ \sum \ln X \ln Y_i &= A \sum \ln X + \beta \sum (\ln X)^2 \end{aligned} \right\} \quad (5.33)$$

نتحصل على جملة معادلة بمجهولين هما  $A$  و  $\beta$ ، حيث:  $A = e^a$  التي يمكن حلها بسهولة بإحدى الطرق الرياضية.

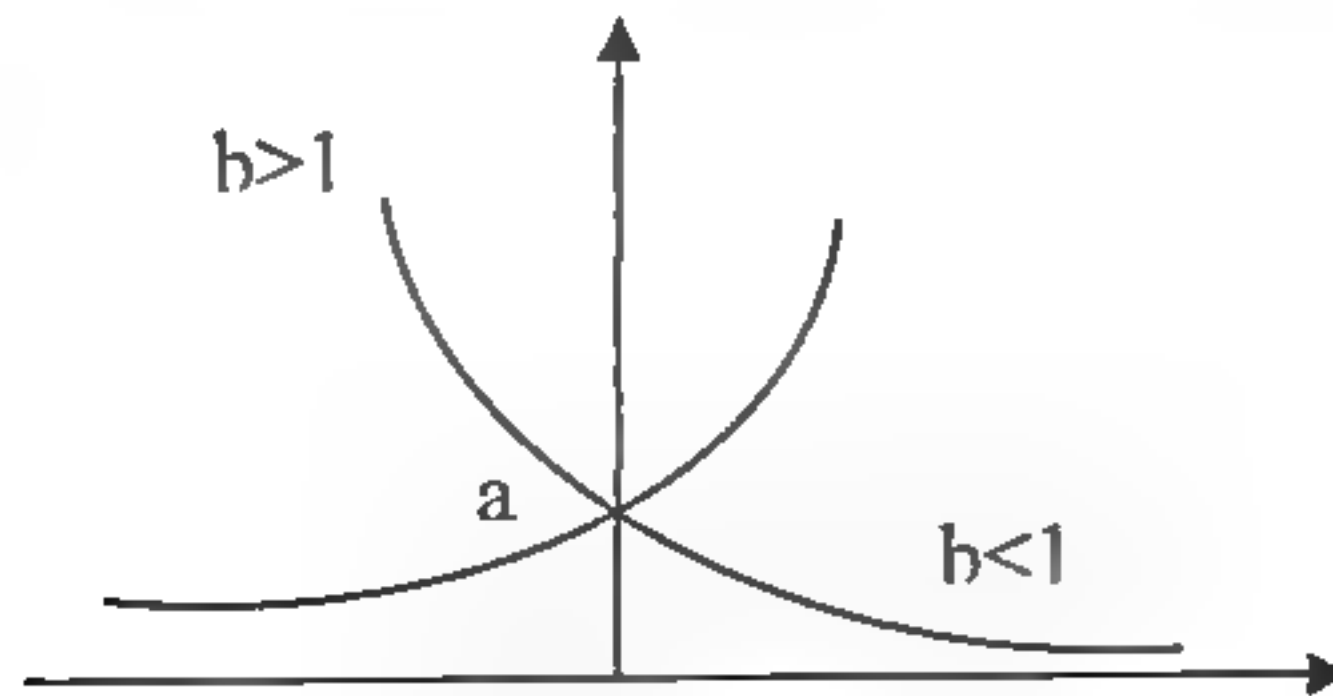
ملاحظة:

يمكن استخدام اللوغارتم العشري في كافة العمليات الحسابية مع التذكير أن:  $A = 10^a$ .

### 5.3.5- التمثيل بواسطة المعادلة الأسية:

إن المعادلات الأسية كثيرة ومتنوعة، وحلها يكون دوما بتحويلها إلى النموذج اللوغارتمي السالف الذكر، ولذلك نكتفي بأخذ الشكل التالي:

- التمثيل بواسطة معادلة من الشكل:  $Y = ab^x$ ، حيث  $b > 0$



التمثيل بواسطة معادلة  $Y = ab^x$



إن منحنيات هذا التمثيل يأخذ شكلين حسب قيمة  $b$  بالنسبة إلى 1

لمعالجة هذه المعادلة، نأخذ لوغاريتم الطرفين فنحصل على:

$$\ln Y = \ln a + X \ln b$$

بفرض أن:  $\ln a = A$  و  $\ln b = B$ ، يصبح لدينا:

$$\ln Y = A + BX$$

لحساب قيم ثوابت هذه المعادلة، نستخدم طريقة المربعات الصغرى

فنحصل على الجملة التالية:

$$\left. \begin{aligned} \sum \ln Y_i &= nA + B \sum X \\ \sum X \ln Y_i &= A \sum X + B \sum X^2 \end{aligned} \right\} (5.34)$$

وذلك بعد الجمع وضرب المعادلة الأصلية في  $x$  ثم الجمع وفق قيم  $i$ . تحل الجملة بإحدى الطرق الرياضية، ومن ثم يتم الرجوع إلى قيم  $a$  و  $b$  بإجراء عمليات تحويل اللوغاريتم (أنظر الطريقة المتبعة في المعادلة (5.31).

### 6.3.5- التمثيل بواسطة المعادلات المثلثية\*:

تستعمل المعادلات المثلثية أساسا في تحليل السلاسل الزمنية، بغية إعطاء نموذج رياضي للتغيرات الدورية و التغيرات الموسمية. و عموما فإن الشكل العام لهذه التغيرات تأخذ النموذج التالي:

$$Y_t^* = a_0 + \sum_{k=1}^m (a_k \sin k\theta_t + b_k \cos k\theta_t) \quad (5.35)$$

$$\theta_t = \frac{X - X_0}{X_n - X_0} 2\pi \quad (5.36) \quad \text{حيث:}$$

---

\* لمزيد من المعلومات، راجع الفصل الرابع الخاص بتحليل السلاسل الزمنية.

ويمكن حساب الثوابت  $a_0$ ،  $a_k$ ،  $b_k$  كما يلي:

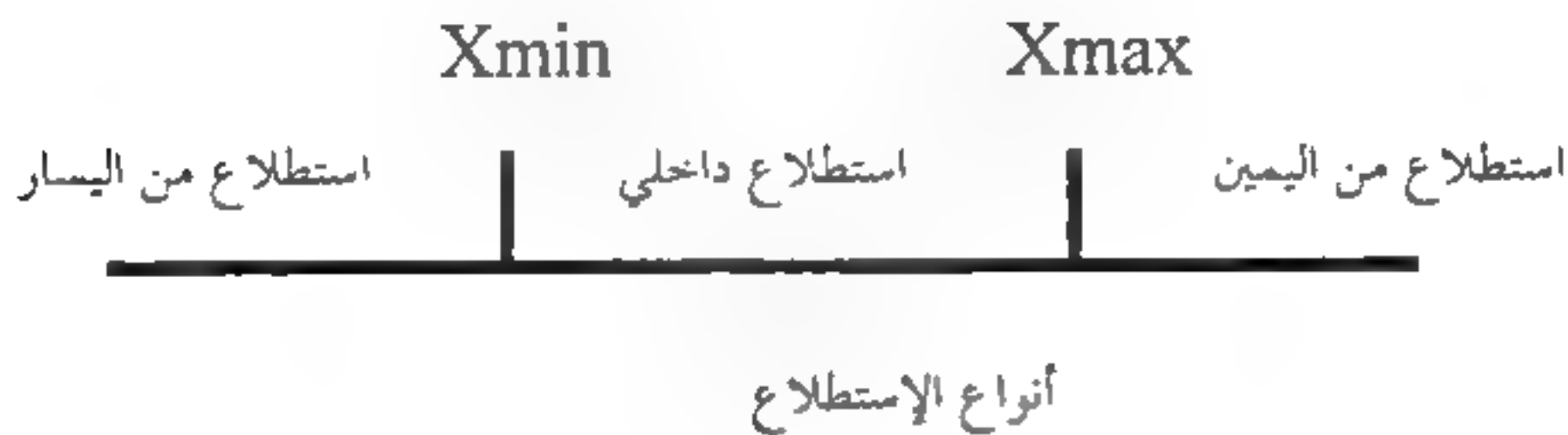
$$\left. \begin{aligned} a_0 &= \frac{1}{n} \sum_{i=1}^n Y_i \\ a_k &= \frac{2}{n} \sum_{i=1}^n Y_i \sin k\theta_i \\ b_k &= \frac{2}{n} \sum_{i=1}^n Y_i \cos k\theta_i \end{aligned} \right\} (5.37)$$

#### 4.5- تطبيقات معادلات التمثيل:

تشكل هذه المعادلات نماذج رياضية يمكن توظيفها في عدة مجالات أهمها:

#### 1.4.5- في مجال الإستطلاع و التنبؤ:

الإستطلاع هو عملية بحث عن قيمة  $Y$  (غير موجودة في السلسلة الارتباطية) بدلالة  $X$  أو العكس، وهو يختلف عن التنبؤ الذي يستخدم في السلاسل الزمنية التي تستخدم متغير الزمن كمتغير مستقل. ويمكن أن نميز ثلاثة أنواع من الإستطلاع وهما: الإستطلاع من الداخل (بأن نعطي قيمة لـ  $X$  تتغير بين  $X_{min}$  و  $X_{max}$ )، الإستطلاع من اليمين (بأن نعطي قيمة لـ  $X$  أكبر من  $X_{max}$ ) والإستطلاع من اليسار (بأن نعطي قيمة لـ  $X$  أصغر من  $X_{min}$ )



## ➤ تمرين:

نعتبر العلاقة الارتباطية التالية:  $C = 16 + \frac{12000}{X}$  ، حيث  $C$  هي كلفة السلعة و  $X$  تمثل كمية إنتاجها، و ذلك في عدد من المؤسسات الإنتاجية، و أن قيم  $X$  تتغير في المجال  $[50,100]$ . فإذا كان تباين خطأ هذا التمثيل يساوي 4 فالمطلوب:

- تقدير كلفة تلك السلعة في مؤسسة جديدة سيبلغ إنتاجها 120 وحدة / اليوم، ثم حساب مجال ثقة هذا التقدير باحتمال قدره 0.96،
- تقدير كلفة تلك السلعة في مؤسسة لم يشملها بحثنا السابق، مع العلم أن إنتاجها يقدر بـ 80 وحدة / اليوم، ثم حساب مجال ثقة هذا التقدير باحتمال قدره 0.997،

## ➤ الحل:

$$\frac{12000}{120} = 116 \quad C_{120} = 16 + 116$$

مجال الثقة عند احتمال 0.96:  $116 - 2.2 \leq C \leq 116 + 2.2$  (أنظر المعادلة 5.14)

ومنه:  $120 \leq C \leq 112$

## 2.4.5- في مجال حساب المؤشرات الاقتصادية المختلفة:

تساهم معادلات التمثيل في جميع أنواع التحليل الإقتصادي، فإذا كانت  $Y$  تمثل حجم الإنتاج في مؤسسة ما و المرتبط بأحد عوامل الإنتاج وليكن  $X$ ، حيث  $Y = f(x)$ ، فإنه يمكن حساب الإنتاجية الحدية وفق العلاقة التالية:

$$\frac{dY}{dX} = \frac{df(X)}{dX} \quad (5.38)$$

كما يمكن كذلك حساب مرونة الإنتاج وفق ما يلي:

$$E = \frac{\frac{dy}{Y}}{\frac{dX}{X}} = \frac{X}{Y} \frac{dY}{dX} \quad (5.39)$$

هذه المعادلة تعبر عن نسبة التزايد النسبي لـ  $Y$  على التزايد النسبي لـ  $X$  عند النقطة  $X$ .

➤ **تمرين:**

لتكن دالة كلفة الإنتاج  $Y$  معرفة بمعادلة التمثيل التالية:

$$Y = X^3 - 10X^2 + 17X + 66$$

**المطلوب:**

- أوجد أصغر قيمة للكلفة إذا كانت  $X$  تتغير في المجال  $[0,10]$
- أوجد حجم الإنتاج الذي يكون فيه الربح أعظميا وذلك بفرض أن سعر الوحدة هو 5 دج
- أوجد مرونة الكلفة بالنسبة لـ  $X$  عند نقطة الربح الأعظمي.

➤ **الحل:**

- أصغر قيمة للكلفة تقابل النهاية الصغرى لدالة الكلفة، ولذلك فإننا نحسب مشتق هذه الدالة:

$$Y' = 3X^2 - 20X + 17$$

$$Y' = 0 \text{ يعني } 3X^2 - 20X + 17 = 0$$

إن حل هذه المعادلة بالمميز يعطي حلين هما:

$$X = 1 \Rightarrow Y(1) = 74 \quad (\text{أكبر قيمة})$$

$$X = 5.67 \Rightarrow Y(5.67) = 22.43 \quad (\text{أصغر قيمة})$$

- لتكن  $P$  هي دالة البيع، حيث  $P = 5X$ ، فتكون دالة البيع:

$$Z = P - Y = 5X - X^3 + 10X^2 - 17X - 66$$

يكون الربح أعظميا إذا كان:

$$Z' = (-X^3 + 10X^2 - 12X - 66)' = 0$$

$$-3X^2 + 20X - 12 = 0$$

بجد حل المعادلة نجد:

$$X = 6 \Rightarrow Z(1) = 6 \quad (\text{ربح أعظمي})$$

$$X = 0.66 \Rightarrow Y(0.66) = -67.63 \quad (\text{خسارة})$$

- مرونة Y عند النقطة  $X = 6$  (عند الربح الأعظمي):

$$E_6 = \frac{X}{Y}(3X^2 - 20X + 17) = \frac{6}{24}(5) = 125$$

### 5.5- الارتباط المتعدد:

لاحظنا في البند الأول أن المتغير المرتبط Y يتأثر بمتغير واحد فقط، يدعى المتغير المستقل، ولذلك سمي بالارتباط البسيط، غير أن الكثير من الظواهر تتأثر بجملة من المتغيرات. نسمى العلاقة بين المتغير Y و جملة المتغيرات  $X_1, X_2, \dots, X_n$  بالارتباط المتعدد.

➤ أمثلة:

- في المجال الزراعي: يتأثر المنتج الزراعي بجملة من العوامل أهمها: الأسمدة  $X_1$ ، طبيعة التربة  $X_2$ ، المناخ  $X_3$ . أي أن  $Y = f(X_1, X_2, X_3)$
- في المجال الاقتصادي: يتأثر مقدار دخل أسرة بمجموعة من المتغيرات أهمها: دخل الأسرة  $X_1$ ، عدد الأفراد  $X_2$  و بمستوى أسعار السلع الاستهلاكية  $X_3$ . أي أن  $Y = f(X_1, X_2, X_3)$ .

### 1.5.5- الخطوات الأساسية لدراسة الارتباط المتعدد:

- تحديد العوامل المؤثرة في المتغير Y (عددا و نوعا)،
- اختيار المتغيرات المستقلة وذلك بحذف أحد أي متغيرين مرتبطين ببعضهما و الإبقاء على المتغيرات المستقلة الهامة،
- حساب معامل الارتباط المتعدد للعلاقة الارتباطية،

- تحديد نوع معادلة التمثيل بناءً على طبيعة العلاقة المدروسة،
- حساب ثوابت معادلة التمثيل بطريقة المربعات الصغرى،
- إختبار معامل التحديد.

#### (أ) - إختيار المتغيرات المستقلة و الهامة:

يتميز الارتباط المتعدد بكثرة متغيراته، و التي إذا دخلت في معادلة التمثيل زادتها تعقيدا، لذلك يحاول الباحثون الإقتصار على الهامة منها فقط في حين يتم حذف باقي المتغيرات. من أجل ذلك فإنه يجب حساب معاملات الارتباط الزوجية بين المتغيرات المستقلة فيما بينها، فيكون:

$$r_{kh} = \frac{\sum (X_{ki} - X_k)(X_{hi} - X_h)}{n\sigma_k\sigma_h} \quad (5.40)$$

حيث أن  $\sigma_{xh}$  و  $\sigma_{xk}$  هما تباينا  $X_h$  و  $X_k$  على الترتيب.

الخطوة التالية تتمثل في حساب معاملات الارتباط الزوجية بين المتغيرات المستقلة والمتغير المرتبط  $Y$  كما يلي:

$$r_{ky} = \frac{\sum (X_{ki} - \bar{X}_k)(Y_i - \bar{Y})}{n\sigma_k\sigma_Y} \quad (5.41)$$

تلخص هذه الحسابات في الجدول التالي الذي يعرف بمصفوفة الارتباط:

المتغيرات	$X_1$	$X_2$	$X_3$	.....	$X_k$	.....	$X_m$	$Y$
$X_1$	1	$r_{21}$	$r_{31}$	.....	$r_{k1}$	.....	$r_{m1}$	$r_1$
$X_2$	$r_{12}$	1	$r_{23}$		$r_{k2}$	.....	$r_{m2}$	$r_2$
$X_3$	$r_{13}$	$r_{23}$	1	.....	$r_{k3}$	.....	$r_{m3}$	$r_3$
				1	1	.....	$r_{mh}$	.....
$X_h$	$r_{1h}$	$r_{2h}$	$r_{3h}$	.....	$r_{kh}$	.....	$r_{km}$	$r_h$
$X_m$	$r_{1m}$	$r_{2m}$	$r_{3m}$	.....	$r_{km}$	.....	1	$r_m$
$Y$	$r_1$	$r_2$	$r_3$	.....	$r_k$	.....	$r_m$	0

## مصفوفة الارتباط

من خلال جدول مصفوفة الارتباط، نقوم بحذف أحد المتغيرين المستقلين اللذان يحققان العلاقة التالية:

$|r_{kh}| \geq 0.80$  حيث  $K \neq h$ ، هذا يعني وجود ارتباط قوي بين هذين المتغيرين. ولكن أي المتغيرين يجب حذفه؟.

لتحديد أي المتغيرين يجب حذفه، نلجأ إلى معاملات الارتباط مع المتغير المرتبط  $Y$ ، فإذا كان:

$|r_{ky}| > |r_{hy}|$  هذا يعني أن المتغير المستقل  $X_k$  له علاقة قوية مع  $Y$ ، و بالتالي لا يمكن حذفه (يحذف عندئذ  $X_h$ ).

➤ مثال تطبيقي:

قمنا بدراسة علاقة المتغير  $Y$  بمجموعة من المتغيرات المستقلة  $X_1, X_2, X_3$ ، فتحصلنا على مصفوفة الارتباط التالية:

المتغيرات	$X_1$	$X_2$	$X_3$	$Y$
$X_1$	1	0.85	0.41	0.71
$X_2$	0.85	1	0.32	0.44
$X_3$	0.41	0.32	1	0.90
$Y$	0.71	0.44	0.9	0

المطلوب: حدد المتغيرات المستقلة الداخلة في معادلة التمثيل.

➤ الجواب:

نلاحظ أن  $|r_{12}| = 0.85 > 0.80$ ، هذا يعني أن  $X_1$  جد مرتبط مع  $X_2$  مما يعني حذف أحدهما. نلاحظ كذلك أن  $r_1 = 0.71$  و  $r_2 = 0.44$  أي أن  $r_1 > r_2$  وبالتالي يحذف  $X_2$ .

(ب) - معامل الارتباط المتعدد:

لنفرض أن  $A$  هي مصفوفة الارتباط التي تم تشكيلها، بعد حذف المتغيرات المستقلة التي لا تدخل في معادلة التمثيل، و هي التي ممثلة في الجدول 1، ولنفرض أن المصفوفة  $B$  هي التي مشكلة في الجدول 1، مع الإشارة إلى أن المصفوفة  $B$  لا تحتوي على السطر الأخير و العمود الأخير الخاصان بالمتغير  $Y$ . فيكون معامل الارتباط المتعدد:

$$r^2 = -\frac{\det A}{\det B} < 0 \text{ المتعدد حيث أن}$$

حيث أن  $\det A$  هو محدد المصفوفة  $A$  و  $\det B$  هو محدد المصفوفة  $B$ .  
بعد التعويض نجد:

$$r^2 = -\frac{\det \begin{vmatrix} 1 & r_{12} & r_{13} & \dots & r_{1m} & r_1 \\ r_{21} & 1 & r_{23} & \dots & r_{2m} & r_2 \\ r_{31} & r_{32} & 1 & \dots & r_{3m} & r_3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ r_{m1} & r_{m2} & r_{m3} & \dots & 1 & r_m \\ r_1 & r_2 & r_3 & \dots & r_m & 0 \end{vmatrix}}{\det \begin{vmatrix} 1 & r_{12} & r_{13} & \dots & r_{1m} \\ r_{21} & 1 & r_{23} & \dots & r_{2m} \\ r_{31} & r_{32} & 1 & \dots & r_{3m} \\ \dots & \dots & \dots & \dots & \dots \\ r_{m1} & r_{m2} & r_{m3} & \dots & r_m \end{vmatrix}} \quad (5.42)$$



يتم حساب معامل الارتباط المتعدد بأخذ الجذر التربيعي لحاصل (5.42)، حيث أن  $-1 \leq r \leq +1$ .

• حالة متغيرين مستقلين:

إذا كان  $X_1$  و  $X_2$  متغيرين عشوائيين و  $Y$  هو المتغير المرتبط، فإن معامل الارتباط المتعدد يكون كما يلي:

$$r^2_{Y(X_1X_2)} = \frac{\det \begin{vmatrix} 1 & r_{12} & r_1 \\ r_{21} & 1 & r_2 \\ r_1 & r_2 & 0 \end{vmatrix}}{\det \begin{vmatrix} 1 & r_{12} \\ r_{21} & 1 \end{vmatrix}} = \frac{-r_2^2 + 2r_{12}r_1r_2 - r_1^2}{1 - r_{12}^2}$$

$$r_{Y(X_1X_2)} = \sqrt{\frac{r_1^2 + r_2^2 - 2r_{12}r_1r_2}{1 - r_{12}^2}} \quad (5.43)$$

➤ تمرين:

يمثل الجدول التالي معطيات إحصائية خاصة بعدد الوحدات الإنتاجية  $Y$  التي ينتجها العامل ومقدار الوقت الضائع خلال فترة العمل  $X_1$  ومدة خبرته المهنية  $X_2$ :

رقم	1	2	3	4	5	6	7	8	9	10
Y	390	387	387	401	394	394	395	391	404	395
$X_1$	19	15	17	11	14	12	16	13	10	13
$X_2$	3	2	3	5	5	4	3	5	6	4

المطلوب: حساب معامل الارتباط المتعدد.

➤ الحل:

$$r_{y(x_1x_2)} = \sqrt{\frac{r_1^2 + r_2^2 - 2r_1r_2r_{12}}{1 - r_{12}^2}}$$

$$r_1 = \frac{\sum (X_{1i} - \bar{X}_1)(Y_i - \bar{Y})}{n\sigma_{x1}\sigma_y} ; r_2 = \frac{\sum (X_{2i} - \bar{X}_2)(Y_i - \bar{Y})}{n\sigma_{x2}\sigma_y} ;$$

$$r_{x1x2} = \frac{\sum (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{n\sigma_{x1}\sigma_{x2}}$$

بعد الحسابات نجد:  $\bar{Y} = 394$  و  $\bar{X}_2 = 4$  ،  $\bar{X}_1 = 14$  حيث أن:

$$r_1 = -0.74 , r_2 = 0.77 , r_{12} = -0.76$$

$$r_{y(x_1x_2)} = \sqrt{\frac{(-0.74)^2 + (0.77)^2 - 2(-0.74)(0.77)(-0.76)}{1 - (-0.76)^2}} = 0.81$$

وهو ما يعني أن علاقة المتغير  $Y$  بالمتغيرين  $X_1$  و  $X_2$  أقوى من علاقته بكل منهما على حدى (لأن  $r_{y(x_1x_2)} > r_1, r_2$ ).

• اختبار دلالة معامل الارتباط المتعدد:

إن هذا الاختبار يخضع إلى حساب الكمية التالية وفق العلاقة:

$$t_r = \frac{r\sqrt{n-m}}{1-r^2} \quad (5.44)$$

حيث:  $n$  هو حجم العينة،  $m$  هو عدد الثوابت الداخلة في معادلة التمثيل. فإذا كان  $t_r > 2$  فإن لمعامل الارتباط المتعدد معنوية.

(ج) - تحديد نوع معادلة التمثيل:

تعتبر هذه المرحلة من أصعب مراحل الارتباط المتعدد، فغالبا ما تكون الصيغة الرياضية المقترحة تعتمد على الأسس الحسية والمادية، كأن ندرس كيفية

تفاعل المتغيرات المستقلة مع المتغير المرتبط، و كيفية تناسبها. كما يجب أن تكون المعادلة المقترحة تحتوي على ثوابت سهلة الحسابات.

## 2.5.5- أهم أنواع المعادلات المستخدمة في الارتباط المتعدد: • الارتباط الخطي المتعدد:

إن المعادلة المقترحة تأخذ الشكل التالي:

$$Y^* = a_0 + a_1X_1 + a_2X_2 + \dots + a_mX_m \quad (5.45)$$

ولحساب ثوابت هذه المعادلة، نستخدم طريقة المربعات الصغرى، فنحصل على المعادلة الأولى:

$$\sum Y = na_0 + a_1 \sum X_1 + a_2 \sum X_2 + \dots + a_m \sum X_m \quad \dots\dots\dots 1$$

نضرب المعادلة الأصلية في  $X_1$  ثم نجمع وفق قيم  $i$  فنحصل على مايلي:

$$\sum X_1 Y = a_0 \sum X_1 + a_1 \sum X_1^2 + a_2 \sum X_1 X_2 + \dots + a_m \sum X_1 X_m \quad \dots\dots\dots 2$$

نضرب المعادلة الأصلية في  $X_2$  ثم نجمع وفق قيم  $i$  فنحصل على مايلي:

$$\sum X_2 Y = a_0 \sum X_2 + a_1 \sum X_2 X_1 + a_2 \sum X_2^2 + \dots + a_m \sum X_2 X_m \quad \dots\dots\dots 3$$

$$\sum X_m Y = a_0 \sum X_m + a_1 \sum X_m X_1 + a_2 \sum X_m X_2 + \dots + a_m \sum X_m^2 \quad \dots\dots\dots n$$

فإذا كان لدينا معادلة بمتغيرين  $X_1$  و  $X_2$  فإننا نحصل على الجملة التالية:

$$\left. \begin{aligned} \sum Y &= na_0 + a_1 \sum X_1 + a_2 \sum X_2 \\ \sum X_1 Y &= a_0 \sum X_1 + a_1 \sum X_1^2 + a_2 \sum X_1 X_2 \\ \sum X_2 Y &= a_0 \sum X_2 + a_1 \sum X_2 X_1 + a_2 \sum X_2^2 \end{aligned} \right\} \quad (5.46)$$

ولحل هذه الجملة، نستخدم طريقة المصفوفات.

➤ مثال تطبيقي:

يمثل الجدول التالي علاقة نفقات أسرة على الأغذية بعدد أفرادها وبمقدار دخلها الشهري:

النفقات على الأغذية Y	عدد الأفراد X <sub>2</sub>	الدخل الشهري X <sub>1</sub>
25	1	90
28	1	110
31	2	120
32	2	130
36	3	180
42	3	220
55	4	280

- نلاحظ أن Y تتناسب طرذا مع X<sub>1</sub> و عكسا مع X<sub>2</sub>، كما أن إنعدام الدخل الشهري لا يعني إنعدام النفقات على الأغذية. وعليه يمكن إقتراح معادلة خطية.

بعد العمليات الحسابية وفق المعادلة (5.46) نحصل على الشكل التالي:

$$\begin{aligned} 7a_0 + 110a_1 + 16a_2 &= 249 \\ 1110a_0 + 202300a_1 + 2960a_2 &= 43490 \\ 16a_0 + 2960a_1 + 44a_2 &= 633 \end{aligned}$$

وبحل هذه الجملة نجد:

$$\begin{aligned} a_0 &= 11.15 \\ a_1 &= 0.16 \\ a_2 &= -0.93 \end{aligned}$$

$$Y^* = 11.15 + 0.16 X_1 - 0.93 X_2 \quad \text{أي أن:}$$

### ملاحظة هامة:

حتى يكون النموذج المقترح يمثل بيانات المثال السابق، يجب التأكد من توافق (تقارب) القيم النظرية، المحسوبة إنطلاقا من معادلة التمثيل، من القيم الحقيقية (وهو ما لوحظ في مثالنا هذا)، أما إذا كان العكس، فإنه يجب التفكير في نموذج آخر مثل النموذج التالي:

#### • الارتباط المتعدد ذو الجداء:

إن النموذج المقترح يأخذ الشكل التالي:

$$Y^* = a_0 X_1^{a_1} \cdot a_1 X_2^{a_2} \cdot \dots \cdot a_{(m-1)} X_m^{a_m} \quad (4.47)$$

حيث أن:  $X_1, X_2, \dots, X_m$  تمثل المتغيرات الداخلة في معادلة التمثيل.

ولحساب ثوابت هذا النموذج، فإنه يجب تحويله إلى نموذج خطي بإدخال اللوغارتم (نيبري أو عشري)، نتحصل على ما يلي:

$$\ln Y^* = \ln a_0 + a_1 \ln X_1 + a_2 \ln X_2 + \dots + a_m \ln X_m$$

فإذا كان لدينا نموذج ذو متغيرين:

$$Y^* = a_0 X_1^{a_1} \cdot a_1 X_2^{a_2}$$

فإن تحويله يعطي المعادلة التالية:

$$\ln Y = \ln a_0 + a_1 \ln X_1 + a_2 \ln X_2$$

ولحساب ثوابت هذه المعادلة، نضع  $A = \ln a_0$  ثم نستخدم الطريقة السابقة بإدخال المجموع على المعادلة:

$$\Sigma \ln Y = nA + a_1 \Sigma \ln X_1 + a_2 \Sigma \ln X_2 \dots \dots \dots 1$$

نضرب المعادلة الأصلية في  $\ln X_1$  ثم نجمع وفق قيم  $i$  فنحصل على مايلي:

$$\sum \ln X_1 \ln Y = A \sum \ln X_1 + a_1 \sum (\ln X_1)^2 + a_2 \sum \ln X_1 \ln X_2 \dots \dots \dots 2$$

نضرب المعادلة الأصلية في  $\ln X_2$  ثم نجمع وفق قيم  $i$  فنحصل على مايلي:

$$\sum \ln X_2 \ln Y = A \sum \ln X_2 + a_1 \sum \ln X_2 \ln X_1 + a_2 \sum (\ln X_2)^2 \dots \dots \dots 3$$

ولحل هذه الجملة، نستخدم طريقة المصفوفات، أما قيمة  $a_0$  فهي:

$$a_0 = e^A \text{ من أجل اللوغارتم النيري}$$
$$a_0 = 10^A \text{ من أجل اللوغارتم العشري}$$

• الارتباط المتعدد الأسّي:

إن هذا الارتباط يأخذ الشكل التالي:

$$Y^* = a_0 e^{a_1 X_1} \cdot a_1 e^{a_2 X_2} \cdot \dots \dots \dots a_m e^{a_{m+1} X_{m+1}} \quad (5.48)$$

حيث  $X_1, X_2, \dots, X_m$  هي المتغيرات المستقلة و الهامة. ولحساب ثوابت هذا النموذج، ندخل اللوغارتم على طرفي المعادلة فنحصل على:

$$\ln Y^* = \ln a_0 + a_1 X_1 + a_2 X_2 + \dots \dots + a_m X_m$$

بوضع  $A_0 = \ln a_0$  ، ومن ثم نستخدم نفس الطريقة السابقة في حالة الارتباط المتعدد ذو الجداء بمتغيرين مستقلين نحصل على:

$$\left. \begin{aligned} \sum \ln Y &= nA_0 + a_1 \sum \ln X_1 + a_2 \sum \ln X_2 \\ \sum (\ln X_1)(\ln Y) &= A_0 \sum \ln X_1 + a_1 \sum (\ln X_1)^2 + a_2 \sum (\ln X_1)(\ln X_2) \dots \dots \dots \\ \sum (\ln X_2)(\ln Y) &= A_0 \sum \ln X_2 + a_1 \sum (\ln X_2) \ln X_1 + a_2 \sum (\ln X_2)^2 \dots \dots \dots \end{aligned} \right\} 5.49$$

هذه الحملة تحل بطريقة المصفوفات كسابقاتها.

• الارتباط المتعدد ذو الجداء المختلط:

إن أشهر مثال على هذا الارتباط هو دالة كوب دوغلاس المشهورة، والتي تستعمل أساسا في الإقتصاد الجزئي. إن هذا الارتباط يأخذ الشكل التالي:

$$Y^* = a_0 X_1^{a_1} \cdot X_2^{a_2} \cdot X_3^{a_3} \cdot e^{a_4 X_4} e^{a_5 X_5} \dots (5.50)$$

ولحساب ثوابت هذه المعادلة، ندخل اللوغارتم على الطرفين ثم نستخدم الطريقة السابقة في تشكيل جملة المعادلات. ولدراسة دالة كوب دوغلاس التي تأخذ الشكل التالي:

$$Q = a_0 K^{a_1} \cdot L^{a_2} \cdot e^{a_3 t}$$

حيث:  $t$ ،  $L$ ،  $K$ ، هما على التوالي الزمن ومقدار العمل و رأس المال. و  $Q$  هي كمية الإنتاج، نأخذ المثال التالي:

يمثل الجدول التالي أرقام حول حجم الإنتاج وكمية رأس المال وكمية العمل لقطاع المحروقات خلال الفترة 1980-2000.

السنة	كمية العمل L	رأس المال K	كمية الإنتاج Q	logL	logK	logQ	logQ.logL	logQ.logK
1980	105	107	101	2,02	2,03	2,00	4,05	4,07
1981	110	114	112	2,04	2,06	2,05	4,18	4,22
1982	118	122	122	2,07	2,09	2,09	4,32	4,35
1983	123	131	124	2,09	2,12	2,09	4,38	4,43
1984	116	138	122	2,06	2,14	2,09	4,31	4,46
1985	125	149	143	2,10	2,17	2,16	4,52	4,68
1986	133	163	152	2,12	2,21	2,18	4,63	4,83
1987	138	176	151	2,14	2,25	2,18	4,66	4,89
1988	121	185	126	2,08	2,27	2,10	4,37	4,76
1989	140	198	155	2,15	2,30	2,19	4,70	5,03
1990	144	208	159	2,16	2,32	2,20	4,75	5,10
1991	145	216	153	2,16	2,33	2,18	4,72	5,10
1992	152	226	177	2,18	2,35	2,25	4,90	5,29
1993	154	236	184	2,19	2,37	2,26	4,95	5,37
1994	149	244	169	2,17	2,39	2,23	4,84	5,32
1995	154	266	189	2,19	2,42	2,28	4,98	5,52
1996	182	298	225	2,26	2,47	2,35	5,32	5,82
1997	196	335	227	2,29	2,53	2,36	5,40	5,95
1998	200	366	223	2,30	2,56	2,35	5,40	6,02
1999	193	387	218	2,29	2,59	2,34	5,34	6,05
2000	193	407	231	2,29	2,61	2,36	5,40	6,17

إن معادلة دالة الإنتاج بدلالة رأس المال و كمية العمل تأخذ الشكل التالي:

$$Y^* = a_0 X_1^{a_1} \cdot X_2^{a_2}$$

حيث:  $X_1$  تمثل رأس المال و  $X_2$  هو كمية العمل. ولحساب ثوابت هذه المعادلة، فإننا نأخذ الجملة التالية:

$$\Sigma \ln Y = nA_0 + a_1 \Sigma \ln X_1 + a_2 \Sigma \ln X_2$$

$$\Sigma (\ln X_1)(\ln Y) = A_0 \Sigma \ln X_1 + a_1 \Sigma (\ln X_1)^2 + a_2 \Sigma (\ln X_1)(\ln X_2)$$

$$\Sigma (\ln X_1)(\ln Y) = A_0 \Sigma \ln X_1 + a_1 \Sigma (\ln X_1)^2 + a_2 \Sigma (\ln X_1)(\ln X_2)$$



حيث:  $A_0 = \ln a_0$

وإعتقادا على حسابات الجدول السابق، فإننا نجد:

$$21A_0 + 48.57a_1 + 45.53a_2 = 46.28$$

$$48.57A_0 + 112.98a_1 + 105.20a_2 = 108.05$$

$$45.35A_0 + 105.20a_1 + 98.09a_2 = 100.151$$

وبعد حل هذه الجملة نجد:

$$A_0 = 0 \Rightarrow a_0 = 1$$

$$a_1 = 0.29$$

$$a_2 = 0.71$$

وتكون دالة الإنتاج كما يلي:  $Q = K^{0.29} \cdot L^{0.71}$

## 6.5- الارتباط الجزئي:

يقصد بالارتباط الجزئي دراسة العلاقة الارتباطية بين المتغير المرتبط  $Y$  وأحد المتغيرات المستقلة، الداخلة في معادلة التمثيل، على أن تعطى قيم ثابتة لباقي المتغيرات المستقلة (غالباً ما يكون المتوسط الحسابي لقيم المتغير الثابت).

➤ مثال:

لنأخذ المعادلة التالية:

$$Y = 11.15 + 0.16X_1 - 0.93X_2$$

يمكن إعطاء قيمة ثابتة لـ  $X_2$  حيث  $X_2 = 2$ ، فتكون المعادلة السابقة كما يلي:

$$Y = 9.29 + 0.16X_1$$

إن هذه المعادلة تعني تأثير المتغير  $X_1$  في  $Y$  بفرض أن  $X_2$  يبقى ثابتاً. إن هذا الارتباط (البسيط) يختلف عن الارتباط بين  $Y$  و  $X_1$  أو بين  $Y$  و  $X_2$  على حدى. هذا يعني أنه إذ لم يكن المتغيران  $X_1$  و  $X_2$  مستقلين عن بعضهما تماماً فإن معادلة الارتباط الجزئي لأي من المتغيرين  $X_1$  أو  $X_2$  المستخرجة من

معادلة الارتباط المتعدد سوف تختلف عن معادلة الارتباط البسيط لأي منهما بقيم الثوابت المختلفة. وذلك لأنه إذا كان  $X_1$  مرتبطا بـ  $X_2$  فإن معادلة الارتباط لأي منهما مع  $Y$  سوف تعكس ضمنا تأثير المتغير الآخر الذي لا تحويه معادلة التمثيل، أما في معادلة الارتباط الجزئي فتعتبره ثابتا.

### 1.6.5- معاملات الارتباط الجزئية:

لدراسة قوة الارتباط الجزئي لـ  $Y$  بأحد المتغيرات المستقلة الداخلة في معادلة التمثيل، نقوم بحساب معاملات الارتباط الجزئية وفق العلاقات التالية:

$$r_{1(2)} = \frac{(r_1 - r_2 \cdot r_{12})}{\sqrt{(1 - r_2^2)(1 - r_{12}^2)}} \quad (5.51)$$

معامل الارتباط الجزئي لـ  $X_1$  بحيث  $X_2$  ثابتا:

$$r_{2(1)} = \frac{(r_2 - r_1 \cdot r_{12})}{\sqrt{(1 - r_1^2)(1 - r_{12}^2)}} \quad (5.52)$$

معامل الارتباط الجزئي لـ  $X_2$ ،  $X_1$  ثابتا

بينما  $r_1$  و  $r_2$  هما معاملات الارتباط الثنائية بين المتغيرين  $Y$  و  $X_1$  و  $X_2$ ، في حين  $r_{12}$  هو معامل الارتباط بين  $X_1$  و  $X_2$ .

مثال: لنفرض معاملات الارتباط التالية:

$$r_1 = 0.91, r_2 = 0.88, r_{12} = 0.96$$

يمكن حساب معاملات الارتباط الجزئية وفق العلاقة (3.46) و (3.47) فنجد:

$$r_{1(2)} = 0.49, \quad r_{2(1)} = 0.05,$$

أما معامل الارتباط المتعدد فيمكن حسابه وفق العلاقة (3.38):

$$r_{Y(X_1X_2)} = \sqrt{\frac{r^2_1 + r^2_2 - 2r_1r_2r_{12}}{1 - r^2_{12}}} = 0.92$$

نلاحظ أن  $r^2(1)$  صغيرة جدا مما يعني أن تأثير  $X_2$  بمعزل عن  $X_1$  يعتبر ضعيفا وهو ما يبدو منطقيا لأن  $r_{21}$  كبير جدا.

## 2.6.5- تطبيقات الارتباط المتعدد في المجال الاقتصادي:

إن معادلة التمثيل ما هي إلا نموذج رياضي يساعد على حل الكثير من المشاكل، من بينها النماذج المستعملة في الإقتصاد. فيمكن حساب:

- المرونة الحدية النسبية:

وتحسب وفق العلاقة التالية:

$$(5.53) \quad \frac{\delta Y}{\delta X_k} \text{ وتمثل السرعة المطلقة لتزايد دالة الإنتاج } Y \text{ بتزايد } X_k$$

- المرونة النسبية:

وهي مرونة دالة الإنتاج بالنسبة للمتغير  $X$  وتعطى بالعلاقة التالية:

$$(5.54) \quad E_k = \frac{\delta Y}{\delta X_k} \cdot \frac{X_k}{Y} \text{ وتمثل السرعة النسبية لتزايد دالة الإنتاج } Y \text{ بتزايد } X_k$$

- المرونة الكلية:

وهي مجموع المرونات النسبية، تعطى بالعلاقة التالية:

$$(5.55) \quad E = \sum_{k=1}^m E_k \text{، حيث } m \text{ عدد المتغيرات الداخلة في معادلة التمثيل.}$$

- المنفعة المتوسطة للمتغير  $X$ :

وتعطى بالعلاقة التالية:

$$(5.56) \quad \frac{Y}{X_k}$$

## تمارين مختارة

### التمرين الأول:

عند دراسة علاقة متوسط إرتفاع أشجار الصنوبر في غابة ما بعمر أشجار تلك الغابة، تحصلنا على المعطيات التالية:

العمر X	20	30	40	50	60	70	80	90	100	110	120
الإرتفاع	9.3	14.1	18.3	21.8	24.2	26.3	28.1	29.6	30.9	32	32.9

- حول ابيانات التالية إلى إرتباط خطي بسيط،
- أحسب معامل الإرتباط ثم إختبر معنويته،
- هل يمكن أن يعطي النموذج الخطي أحسن تفسير لهذه العلاقة؟، ماذا تقترح؟،
- لنفترض أن النموذج اللوغارتمي هو الذي يعطي التفسير الأحسن لطبيعة هذه العلاقة، بين كيف ذلك،
- شكل هذا النموذج، أحسب معامل تحديده ثم إختبر معنويته.

### التمرين الثاني:

لنفترض أنه لدينا المعطيات التالية عن ستة مؤسسات إنتاجية:

المؤسسة	1	2	3	4	5	6
كمية الإنتاج بالآلاف القطع	2	3.5	4	4.5	5.5	6
قيمة الأصول الثابتة بالمليون	1.9	1.7	1.8	1.6	1.5	1.4

### المطلوب:

- إختيار المعادلة الملائمة لتمثيل هذه العلاقة،
- حساب القيم النظرية لكمية الإنتاج،
- حساب معامل التحديد و إختبار موضوعيته،
- أرسم الخط البياني لهذه العلاقة الإرتباطية.

### التمرين الثالث:

إذا كانت 100 درجة حرارة مئوية تقابل 212 درجة مقدرة بالفهرنهايت، في حين أن درجة الصفر في المئوية تقابلها 32 في وحدة الفهرنهايت. على افتراض وجود علاقة خطية بين وحدة الدرجة المئوية وحدة الفهرنهايت، أوجد:

- المعادلة التي تربط بين الوحدتين (افترض أن المئوية تمثل  $X$ )
- ما هي القيمة التي تقابل 80 درجة مئوية،
- ما هي القيمة التي تقابل 68 فهرنهايت.

### التمرين الرابع:

عينة حجمها 18 ومعامل إرتباطها 0.32. هل يمكن أن نستنتج عند مستوى المعنوية 5% و 1% أن معامل الإرتباط في المجتمع يختلف عن الصفر؟. أوجد أصغر قيمة لـ  $n$  حتى يكون معامل إرتباط العينة السابق يختلف معنوياً عن الصفر ( $\alpha = 5\%$ ).

### التمرين الخامس:

لتكن لدينا البيانات التالية:

$X_1$	64	71	53	67	55	58	77	57	56	51	76	68
$X_2$	57	59	49	62	51	50	55	48	52	42	61	57
$X_3$	8	10	6	11	8	7	10	9	10	6	12	9

### المطلوب:

- أوجد العلاقة الارتباطية التي تجمع بين  $X_1$  كمتغير مرتبط و  $X_2, X_3$ .
- أحسب عندئذ معاملات الارتباط الزوجية المقابلة،
- أحسب قيمة معامل الارتباط المتعدد،
- أحسب معامل التحديد ثم أختبر معنويته،
- قدر قيمة  $X_1$  عند  $X_2 = 54$  و  $X_3 = 9$ .

## الفصل السادس

### السلاسل الزمنية و تحليلها

#### مقدمة:

إن الكثير من الظواهر، خاصة الاجتماعية و الاقتصادية، تتغير بتغير الزمن الذي هو بحد ذاته متغير موضوعي مستقل. وبما أن الزمن ملازم للظواهر فيمكننا أن نربط بين حالة أية ظاهرة و بين اللحظة التي تقابل هذه الحالة أو بين مجمل تطورات ظاهرة و بين الفترة الزمنية التي جرت فيها تلك التطورات.

#### 1.6- تعريف السلسلة الزمنية:

هي عبارة عن سلسلة من القيم العددية لمؤشر إحصائي يعكس تغير ظاهرة ما بالنسبة للزمن، بحيث أن لكل قيمة لإحصائية فترة زمنية تقابلها. يكون متغير الزمن  $t$  متغيرا مستقلا تقابلها قيمة إحصائية مرتبطة  $Y_t$ .

#### 2.6 - نواع السلاسل الزمنية:

يمكن أن نميز نوعين من السلاسل الزمنية:

- السلاسل الزمنية الأنية: وهي السلاسل الزمنية التي حدودها  $Y_t$  تقابل لحظات زمنية معينة. مثل عدد سكان مدينة ما في أحد أيام السنة.

- السلاسل الزمنية المديدة: وهي السلاسل الزمنية التي حدودها  $Y_t$  تقابل فترات زمنية معينة. مثل كمية الإنتاج خلال شهر أو عدد المواليد خلال سنة.

#### 3.6- مؤشرات السلسلة الزمنية:

يستخدم في تحليل السلاسل الزمنية عدد من المؤشرات الإحصائية التي تحدد مقدار واتجاه و سرعة تغير الظاهرة المدروسة عبر الزمن، وهي:

(أ) - الزيادة المطلقة: مقدار زيادة أو نقصان حدود السلسلة الزمنية خلال فترة زمنية معينة، و تعطى بالعلاقة التالية:

$$\Delta Y_{tk} = Y_t - Y_{t-k} \quad (6.1) \quad .$$

حيث  $Y_t$  هو حد السلسلة التي تقابل الفترة  $t$  و  $Y_{t-k}$  حد السلسلة التي تقابل الفترة  $t-k$ ، يسمى هذا الحد بالحد الأساسي. إذا كان  $k=1$  فإن:

$$\Delta Y_{t1} = Y_t - Y_{t-1}$$

➤ مثال:

يمثل الجدول التالي إنتاج الكهرباء في محطة توليد:

السنة	كمية الإنتاج (ألف كيلواط / الساعة)	التزايدات المطلقة $\Delta Y_{t1} = Y_t - Y_{t-1}$	التزايدات المطلقة بالنسبة لعام 1970 $\Delta Y_{t70}$
1970	741	800-741=59	800-741=59
1971	800	857-800=57	857-741=116
1972	857	915-857=58	915-741=174
1973	915		

كيفية حساب الزيادة المطلقة في سلسلة زمنية.

(ب) - معدل النمو: و يعطى بالعلاقة التالية:

$$T_{tk} = \frac{Y_t}{Y_{t-k}} 100 \quad (6.2) \quad , \text{ حيث } Y_{t-k} \text{ يمثل الحد الأساسي،}$$

$$T_{t1} = \frac{Y_t}{Y_{t-1}} 100 \quad \text{إذا كان: } K=1 \text{ فإن:}$$

نسمي الكميات  $T_{t1}$  بمعدلات النمو المتتالية.

- إذا كانت  $T_{t1} > 100$  فإن الحدود تكون متزايدة،

- إذا كانت  $T_{t1} = 100$  فإن الحدود تكون ثابتة،

- إذا كانت  $T_{t1} < 100$  فإن الحدود تكون متناقصة.



في المثال السابق (جدول 1.6) نجد معدلات نمو الإنتاج المتتالية:

$$T_{2i} = \frac{800}{741} \cdot 100 = 107.96\% ; \quad T_{3i} = \frac{857}{800} \cdot 100 = 107.13\% ;$$

$$T_{2i} = \frac{915}{857} \cdot 100 = 106.77\%.$$

(ج) - معدل الزيادة النسبية: ويعرف بالعلاقة التالية:

$$m_{tk} = \frac{Y_t - Y_{t-k}}{Y_{t-k}} 100 \quad (6.3)$$

إذا كان:  $K=1$  فإن  $m_{1i} = \frac{Y_t - Y_{t-1}}{Y_{t-1}} 100$ ، وهو يمثل مقدار تزايد النسبة

المثوية للحد  $Y_t$  بالنسبة للأساس  $Y_{t-k}$

(د) - المتوسط الحسابي لحدود سلسلة زمنية: و يعطى بالعلاقة التالية:

$$\bar{Y} = \frac{\sum_{t=1}^n Y_t}{n} \quad (6.4)$$

(هـ) - المتوسط التوافقي لحدود سلسلة زمنية: يعطى بالعلاقة التالية:

$$\bar{Y}_H = \frac{1/2 Y_1 + Y_2 + \dots + Y_{n-1} + 1/2 Y_n}{n-1} \quad (6.5)$$

(و) - متوسط الزيادة المطلقة: يعطى بالعلاقة التالية:

$$\Delta \bar{Y} = \frac{Y_n - Y_1}{n-1} \quad (6.6)$$

(ي) - متوسط معدل النمو: ويعطى بالعلاقة التالية:

$$\bar{T} = \sqrt[n-1]{\frac{Y_n}{Y_1}} \cdot 100 \quad (6.7)$$

ك) - متوسط معدل الزيادة النسبية:

$$\bar{m} = (\bar{T} - 100) \% \quad (6.8)$$

#### 4.6- تحليل السلسلة الزمنية:

يبدأ تحليل الزمنية على إيجاد المسار العام لتطور الظاهرة بدلالة الزمن، وذلك من خلال رسم شكل الانتشار بأخذ عامل الزمن كمتغير مستقل، في حين تشكل قيم الظاهرة المدروسة المتغير المرتبط. ونسمي توصيل نقاط الانتشار فيما بينها بالمنحنى التاريخي للسلسلة الزمنية.

#### 1.4.6- تغيرات الاتجاه العام:

وهي التغيرات التي تعكس مسار تطور الظاهرة المدروسة عبر الزمن، وهي تعطي فكرة واضحة عن تزايد أو تناقص السلسلة الزمنية، بغض النظر عن جميع الانحرافات أو التقلبات.

تشبه معادلة الاتجاه العام تلك التي تم تقديمها في نظرية الارتباط (الفصل الثالث)، يكون فيها عامل الزمن هو المتغير المستقل. وتأخذ، في الغالب، أحد الأشكال التالية:

$$Y^* = a_0 + a_1 t$$

$$Y^* = a_0 + a_1 t + a_2 t^2$$

$$Y^* = a_0 + a_1 t + a_2 t^2 + a_3 t^3$$

$$Y^* = a_0 + a_1 \ln t$$

$$Y^* = a_0 + \frac{a_1}{t} ; \quad Y^* = a_0 t^{a_1}$$

إن عملية اختيار معادلة التمثيل عملية صعبة، لكون بيانات السلسلة الزمنية متذبذبة، ولذلك فلا بد من تسوية هذا التذبذب.

## • المتوسطات المتحركة:

لعل من أهم طرق تسوية تذبذب السلسلة الزمنية هي المتوسطات المتحركة. لنفرض أنه لدينا السلسلة الزمنية التالية:

$$\begin{array}{l} Y: y_1 \quad y_2 \quad y_3 \dots y_t \dots y_n \\ t: t_1 \quad t_2 \quad t_3 \dots t_t \dots t_n \end{array}$$

إن المتوسط المتحرك بطول 3 يكون وفق العلاقة التالية:

$$\begin{aligned} \bar{Y}_2 &= \frac{Y_1 + Y_2 + Y_3}{3} \quad \text{يقابل اللحظة الزمنية: } \bar{t}_2 = \frac{t_1 + t_2 + t_3}{3} \\ \bar{Y}_3 &= \frac{Y_2 + Y_3 + Y_4}{3} \quad \text{يقابل اللحظة الزمنية: } \bar{t}_3 = \frac{t_2 + t_3 + t_4}{3} \\ \bar{Y}_{n-1} &= \frac{Y_{n-2} + Y_{n-1} + Y_n}{3} \quad \text{يقابل اللحظة الزمنية: } \bar{t}_{n-1} = \frac{t_{n-2} + t_{n-1} + t_n}{3} \end{aligned}$$

وهكذا نحصل في الأخير على سلسلة زمنية جديدة أقل تذبذبا من الأولى، وحدودها هي المتوسطات المتحركة المحسوبة سابقا مع أزمنتها.

وبصفة عامة، فإن عدد حدود المتوسطات المتحركة يكون مساويا إلى طور التغيرات الدورية (عدد التواءات في المنحنى التاريخي)، فإذا كان طور التغيرات هو  $P$  فإن علاقة المتوسط المتحرك بأخذ العلاقة التالية:

$$\bar{Y}_{j+\frac{p}{2}} = \frac{\sum_{t=j}^{j+p-1} Y_t}{p} \quad \text{حيث } j=1,2,\dots,n-p \quad (6.9)$$

## ➤ مثال:

لنفترض أنه لدينا السلسلة الزمنية التالية:

السنة	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
إنتاج الغاز (م،م3)	9,5	8,2	8,2	5,9	9,9	5,5	12	8,9	12,2	10

نلاحظ بعد رسم المنحنى التاريخي وجود عدد كبير من التواءات والتذبذبات، لذلك يمكننا التقليل منها باستخدام المتوسطات المتحركة (بطول 3). بعد الحسابات نجد مايلي:

السنة	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
إنتاج الغاز (م،م3)	9,5	8,2	8,2	5,9	9,9	5,5	12	8,9	12,2	10
المتوسطات المتحركة		8,63	7,43	8,00	7,10	9,13	8,80	11,03	10,37	

من خلال المنحنى (الشكل السابق) نلاحظ أن شكل إنتشار السلسلة بدا أكثر وضوحا من شكل الإنتشار الأساسي، مما يساعدنا على إختيار نوع معادلة التمثيل. ولإيجاد قيم ثوابت معادلة التمثيل، فإننا نعتمد على طريقة المربعات الصغرى، ونستخدم نفس الأسلوب المتبع في دراسة الارتباط البسيط والارتباط المنحني\*.

في المثال السابق، يمكن إيجاد معادلة الاتجاه العام، بفرض أن السلسلة الزمنية تأخذ معادلة مستقيم، فنجد:

$$Y = 0,3024t + 7,3667$$

(باستعمال البيانات الأولية) حيث أن:  $R^2 = 0,16$

$$Y = 0,4184t + 6,968$$

(باستعمال بيانات المتوسطات المتحركة) حيث أن:  $R^2 = 0,60$

\*يمكن للقارئ أن يعود إلى هذه المعادلات التي تعرضا إليها بالتفصيل في الفصل الثالث (نظرية الارتباط)

## 2.4.6- التغيرات الدورية:

وهي التغيرات الناجمة عن تأثير القوى الدورية و التي تظهر دوريا من حين لآخر، ويظهر تأثيرها على قيم السلسلة الزمنية على شكل تزايد لهذه القيم (أو تناقص) حتى تبلغ قيمة عظمى (أو صغرى) ن أي على شكل نتوءات.

إن دراسة وإظهار التغيرات الدورية في تطور السلسلة الزمنية يمكن أن يكون بطريقتين:

### • طريقة الانحرافات النظرية:

وتعتمد على حساب إنحرافات القيم النظرية، المحسوبة من خلال معادلة التمثيل، عن القيم الحقيقية، أي أن:  $Z_t = Y_t - Y_t^*$  ومن ثم تمثيل هذه الانحرافات بواسطة معادلة تمثيل تأخذ شكلا جيبيا.

### • طريقة المتوسطات المتحركة:

وتعتمد على حساب المتوسطات المتحركة وفق المعادلة (6.9). فتكون قيم  $Z$  كما يلي:

$$Z_j + \frac{p}{2} = Y_j + \frac{p}{2} - Y^*_{j + \frac{p}{2}} \quad (6.10)$$

حيث أن:  $(j = 1, 2, \dots, n-p)$  و  $Y_{j + \frac{p}{2}}$  هي القيم الحقيقية التي تقابل اللحظة  $t = j + \frac{p}{2}$ .

إن معادلة تمثيل التغيرات الدورية تأخذ الصيغة التالية:

$$Y_t = a_0 + f(\sin \theta_t, \cos \theta_t) \quad (6.11)$$

الخطوات: لنفترض السلسلة الزمنية التالية:

$$\begin{array}{l} Y: y_1 \ y_2 \ y_3 \dots Y_t \dots y_n \\ t: t_1 \ t_2 \ t_3 \dots t_t \dots t_n \end{array}$$

يمكن إجراء تحويل المتغير  $t$  إلى المتغير  $\theta$  وذلك وفق المعادلة التالية:

$$\theta_t = \frac{t - t_0}{t_n - t_0} 2\pi \quad (6.12)$$

تتوصل على سلسلة جديدة بدلالة  $\theta$  حيث:

$$\begin{array}{l} Y: y_1 \ y_2 \ y_3 \dots Y_t \dots y_n \\ \theta: \theta_1 \ \theta_2 \ \theta_3 \dots \theta_t \dots \theta_n \end{array}$$

فإذا كانت السلسلة الزمنية مكونة من  $m$  دورة فإن معادلة تمثيل التغيرات الدورية تعطى بالعلاقة التالية:

$$Y_t^* = a_0 + \sum_{k=1}^m (a_k \sin k\theta_t + b_k \cos k\theta_t) \quad (6.13)$$

أما ثوابت هذه المعادلة فيمكن حسابهم من العلاقات التالية:

$$\left. \begin{array}{l} a_0 = \frac{1}{n} \sum_{i=1}^n Y_i \\ a_k = \frac{2}{n} \sum_{i=1}^n Y_i \sin k\theta_i \\ b_k = \frac{2}{n} \sum_{i=1}^n Y_i \cos k\theta_i \\ k = 1, 2, \dots, m \end{array} \right\} \quad (6.14)$$

الخطوة الأخيرة، بعد إيجاد معادلة تمثيل التغيرات الدورية، تتمثل في إضافة هذه المعادلة إلى المعادلة العامة للسلسلة الزمنية، أي أن:

$$\hat{Y}_t = \tilde{Y}_t + Y_t^* \quad (6.15)$$

مع الإشارة إلى أن:  $\tilde{Y}$  هي القيمة الحقيقية للسلسلة الزمنية،  $\tilde{Y}$  هي قيمة الاتجاه و  $Y^*$  هي قيمة التغيرات الدورية.

➤ مثال:

يمكننا حساب قيم التغيرات الدورية من المثال السابق:

السنة	1990	1991	1992	1993	1994	1995	1996	1997	1998	1971	المجموع
t	1	2	3	4	5	6	7	8	9	10	
إنتاج الغاز Y(م,م3)	9,5	8,2	8,2	5,9	9,9	5,5	12	8,9	12,2	10	90,3
θ	0	40	80	120	160	200	240	280	320	360	
2θ	0	80	160	240	320	400	480	560	640	720	
sinθ	0	0,64	0,98	0,87	0,34	-0,34	-0,87	-0,98	-0,64	0	
Ysinθ	0,00	5,27	8,08	5,11	3,39	-1,88	-10,39	-8,76	-7,84	0	-7,04
sin2θ	0,00	0,98	0,34	-0,87	-0,64	0,64	0,87	-0,34	-0,98	0	
Ysin2θ	0	8,08	2,80	-5,11	-6,36	3,54	10,39	-3,04	-12,01	0	-1,72
cosθ	1	0,77	0,17	-0,50	-0,94	-0,94	-0,50	0,17	0,77	1	
cos2θ	1	0,17	-0,94	-0,50	0,77	0,77	-0,50	-0,94	0,17	1	
Ycosθ	9,5	6,28	1,42	-2,95	-9,30	-5,17	-6,00	1,55	9,35	10	14,68
Ycos2θ	9,5	1,42	-7,71	-2,95	7,58	4,21	-6,00	8,36	2,12	10	9,82

فإذا فرضنا أن السلسلة السابقة مكونة من 2 دورات (أي أن  $m = 2$ )، فإن معادلة تمثيل التغيرات الدورية تعطى بالعلاقة التالية:

$$Y^* = a_0 + a_1 \sin \theta_t + b_1 \cos \theta_t + a_2 \sin 2 \theta_t + b_2 \cos 2 \theta_t$$

نقوم بحساب ثوابت هذه المعادلة من خلال العلاقة (6.14)، فنحصل على

مايلي:

$$= \frac{1}{10} 90.3 = 9.03 \quad a_0 = \frac{1}{n} \sum_{i=1}^n Y_i$$

$$a_k = \frac{2}{n} \sum_{i=1}^n Y_i \sin k\theta_i$$

$$a_2 = \frac{2}{10} (-1.72) = -0.34 \quad , \quad a_1 = \frac{2}{10} (-7.04) = -1.40$$

$$, \quad b_k = \frac{2}{n} \sum_{i=1}^n Y_i \cos k\theta_i$$

$$, \quad b_2 = \frac{2}{10} 9.82 = 1.96 \quad b_1 = \frac{2}{10} 14.68 = 2.90$$

نتحصل بعد التعويضات على المعادلة التالية:

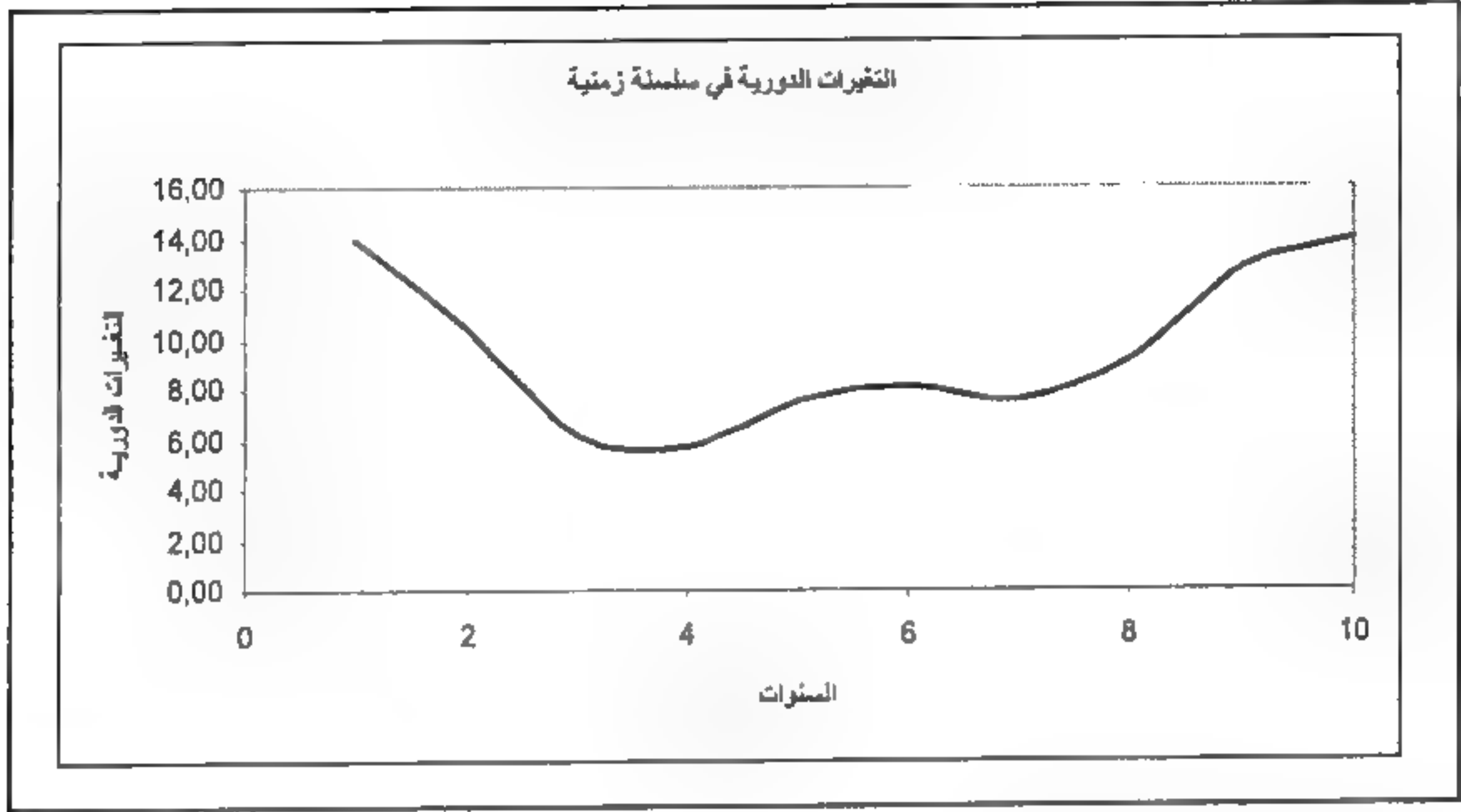
$$Y^* = 9.03 - 1.40 \sin \theta_t + 2.90 \cos \theta_t - 0.34 \sin 2 \theta_t + 1.96 \cos 2 \theta_t$$

نلخص قيم التغيرات الدورية بعد تعويض قيم  $\theta$  في المعادلة فنحصل على الجدول التالي\*:

Y*	13,89	10,36	6,20	5,68	7,55	8,07	7,52	9,19	12,83	13.89
----	-------	-------	------	------	------	------	------	------	-------	-------

\* في مثالنا السابق، كان من المفروض أن يكون عدد الدورات مساويا لـ 4، أي بعدد التواءات، ولتسهيل الحسابات، افترضنا وجود دورتين، فيمكن للقارئ الحصول على منحنى بأربع دورات إذا أخذ  $m = 4$





### 3.4.6- التغيرات الموسمية:

وهي التغيرات الناتجة عن تأثير الفصول خلال السنة أو تأثير الأشهر خلال الفصول أو تأثير الأيام خلال الأشهر. إن دراسة التغيرات الموسمية تشبه إلى حد كبير التغيرات الدورية، فما هي في الحقيقة إلا تغيرات دورية ولكنها مرتبطة بفصول السنة أو الأشهر. فمثلا إرتفاع كمية إستهلاك الغاز خلال موسم الشتاء، ولعدة سنوات متتالية فإذا فرضنا السنة تتكون من 4 فصول، فإنه يمكن تكرار ما جاء في دراسة التغيرات الدورية مع الإشارة أن قيمة  $m$  يجب أن تأخذ عدد المواسم.

### 4.4.6- التغيرات العشوائية:

وهي التغيرات الناجمة عن تأثير قوى عرضية كالزلازل والحروب والإضطرابات العمالية... الخ، ولذلك يمكن إرجاءه إلى خطأ مرتكب يمكن تقديره. وعليه تكون معادلة تمثيل السلسلة الزمنية كما يلي:

$$f(t) = \tilde{Y}(t) + Y^*(t) + \varepsilon(t) \quad (6.15)$$

حيث  $\tilde{Y}(t)$ ,  $Y^*(t)$ ,  $\varepsilon(t)$  هي على التوالي: معادلة الاتجاه العام، معادلة التغيرات الدورية والتغيرات الموسمية معا ومعادلة التغيرات العرضية. ونظرا لصعوبة تقدير التغيرات العرضية وصعوبة تمثيلها رياضيا لكونها عرضية، فإنه

يستحسن تقديرها بمعرفة مركبات المعادلة (6.15) على أساس أن  $f(t)$  هي القيمة الحقيقية للسلسلة الزمنية وأن  $\tilde{Y}(t)$  و  $Y^*(t)$  تم حسابهم آنفا.

### 5.6-التنبؤ:

التنبؤ هو إعطاء قيمة زمنية في معادلة تمثيل الاتجاه العام، غير موجودة في السلسلة، من أجل الحصول على قيمة سلسلة زمنية مقابلة لهذه اللحظة. وينقسم إلى:

#### 1.5.6- التنبؤ الداخلي:

هو إعطاء قيمة غير موجودة في السلسلة الزمنية ومقابلة للحظات الزمنية التي تقع داخل المجال الزمني للسلسلة،

#### 2.5.6- التنبؤ الخارجي:

ويهدف إلى إيجاد قيم مجهولة للسلسلة مقابلة للحظات الزمنية التي تقع خارج المجال الزمني للسلسلة.

#### 3.5.6- طرق التنبؤ:

##### 1.3.5.6- طريقة التمديد:

وتعتبر أسهل طرق التنبؤ، وتعتمد على معادلة الاتجاه العام من خلال إعطاء قيم للزمن في معادلة التمثيل. وينجم عن عملية التنبؤ كمية من الخطأ يمكن تقديرها بحساب تباين الخطأ  $S_{yy}^*$ ، ولذلك يمكن حساب كذلك مجال ثقة هذا التنبؤ باحتمال معلوم حيث:

$$\tilde{Y} - ZS_{yy}^*, \tilde{Y} + ZS_{yy}^* [ ] \quad (6.16)$$

حيث أن  $\tilde{Y}$  هي القيمة التقديرية للسلسلة الزمنية و  $Z$  هو الإحتمال المعلوم. غير أن هذه الطريقة لا تصلح للتنبؤات الطويلة الأمد لأنها تفترض أن الخطأ المرتكب يبقى ثابتاً، وهذا ما يتنافى مع الواقع، إذ أن الخطأ يزداد كلما إزداد مدى التنبؤ

### 2.3.5.6- طريقة الخطأ المتزايد:

في هذه الطريقة يرتبط الخطأ المرتكب بعدد حدود السلسلة الزمنية  $n$  وبمدى التنبؤ  $I$ ، ويكون الخطأ النسبي:

$$\varepsilon_{pr} = \left| \frac{Y_{n+I} - Y'_{n+I}}{Y_{n+I}} \right| \cdot 100 \quad (6.17)$$

حيث أن:  $Y_{n+I}$  هي القيمة الحقيقية التي نريد التنبؤ عنها والتي تقابل اللحظة  $n+I$  و  $Y'_{n+I}$  هي القيمة التقديرية المتنبأ بها. فإذا كان  $L$  هو عدد التنبؤات، فإن متوسط الخطأ النسبي هو:

$$\bar{\varepsilon}_{pr} = \frac{1}{L} \sum_{I=1}^L \left| \frac{Y_{n+I} - Y'_{n+I}}{Y_{n+I}} \right| \cdot 100 \quad (6.18)$$

وحيث أن  $Y_{n+I}$  تبقى دوماً مجهولة فإن يتعذر حساب مقدار  $\bar{\varepsilon}_{pr}$ . ولذلك نلجأ إلى طريقة تطبيقية تمكننا من حساب المعادلة (6.18)، و تلخص فيما يلي:

لنفترض أنه لدينا السلسلة الزمنية التالية:

$$\begin{array}{l} Y: y_1 \quad y_2 \quad y_3 \dots Y_t \dots y_n \\ t: 1 \quad 2 \quad 3 \dots \dots \dots n \end{array}$$

1 - تقسم السلسلة الزمنية إلى قسمين حسب قيمة  $n$ . ويمكن أن نميز:

إذا كان  $n$  زوجي، فإن السلسلة تقسم إلى  $\frac{n}{2}$  حداً. أما إذا كان  $n$  فردي فإن السلسلة تقسم إلى  $\frac{n+1}{2}$  حداً و  $\frac{n-1}{2}$  حد. نأخذ القسم الأول ونعتبره سلسلة زمنية مستقل عن القسم الثاني،

2- نقوم بإيجاد معادلة التمثيل للقسم الأول على أن نتنبأ بقيم القسم الثاني لنتحصل على قيمه النظرية والتي تختلف عن قيم القسم الأول (قيم حقيقية)،

2- نقوم بحساب الأخطاء النسبية التي من شأنها ترتكب من جراء عملية التنبؤ وفق العلاقة (6.18) ومن ثم نحسب متوسط هذه الأخطاء وفق العلاقة (6.19)، حيث أن  $L$  هو عدد الحدود المتنبأ بها،

- 4- نقوم بتسجيل قيم  $\bar{\epsilon}_{pr}$  التي توافق عدد الحدود  $L$ ،
- 5- نضيف إلى حدود القسم الأول حداً إضافياً من القسم الثاني فنحصل من جديد على سلسلتين، الأولى بها  $(n_L+1)$  حداً، و الثانية على  $(n-(n_L+1))$  حداً،
- 6- نكرر الخطوات 2، 3، 4، 5 حتى الوصول إلى معظم حدود القسم الثاني من السلسلة. نشكل معادلة التمثيل التالية:  $\tilde{\epsilon} = f(n, L)$ ، وهي تقدير لقيم متوسط الأخطاء النسبية بدلالة  $n$  و  $L$ .

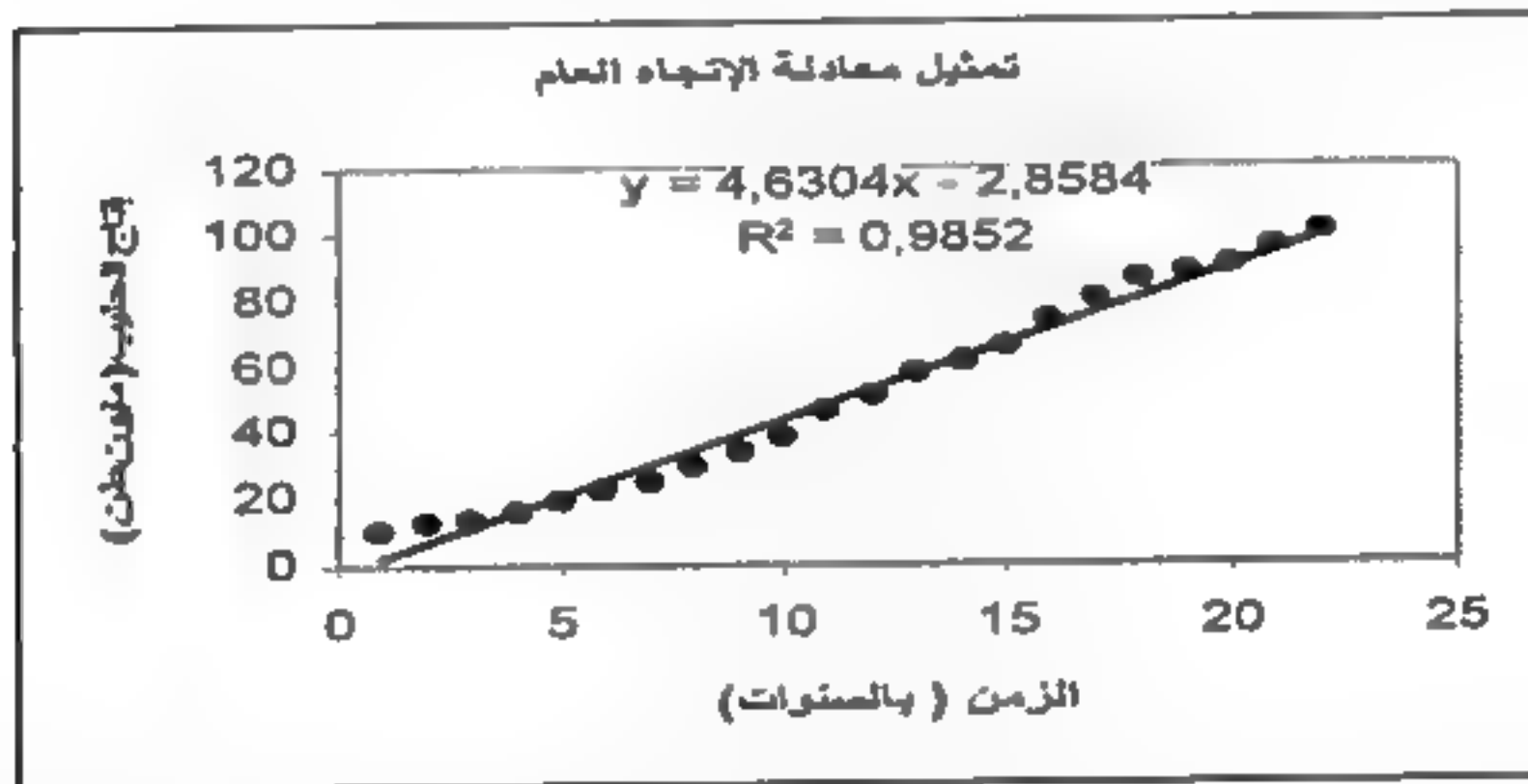
➤ مثال:

يمثل الجدول 2 بيانات سلسلة زمنية عن إنتاج الحليب في منطقة ما. نريد أن نتبأ على هذا الإنتاج إلى ما بعد 2001.

يمكن إيجاد معادلة الاتجاه العام بطريقة المربعات الصغرى، حيث أن  $t$  هو المتغير المستقل و إنتاج الحليب هو المتغير المرتبط، فنجد:

$$Y^* = -2.85 + 4.53t$$

حيث أن  $R^2 = 0.98$ ، وهو ما يشير إلى جودة هذا التمثيل.



وقبل حساب هذا التنبأ، يجب معالجة الأخطاء المرتكبة عنه. نتائج هذه المعالجة مبينة في الجدول التالي الذي يمثل نتائج طريقة الخطأ المتزايد:

## جدول انتاج الحليب لمنطقة ما من سنة 1980 إلى سنة 2001

السنة	t	انتاج الحليب (مطن)	الحالة الأولى $\epsilon_{pr}$	الحالة الثانية $\epsilon_{pr}$	الحالة الثالثة $\epsilon_{pr}$	الحالة الرابعة $\epsilon_{pr}$	الحالة الخامسة $\epsilon_{pr}$	الحالة السادسة $\epsilon_{pr}$	الحالة السابعة $\epsilon_{pr}$
1980	1	10,2	$Y^* = 3.74 + 3.40t$	$Y^* = 2.67 + 3.53t$	$Y^* = 1.51 + 3.88t$	$Y^* = 0.81 + 4.02t$	$Y^* = 0.32 + 4.11t$	$Y^* = -0.44 + 4.25t$	$Y^* = -1.4 + 4.41t$
1981	2	12,1							
1982	3	13,9							
1983	4	16							
1984	5	19							
1985	6	22,5							
1986	7	24,9							
1987	8	28,9							
1988	9	33,3							
1989	10	38,3							
1990	11	45,5							
1991	12	50,9	12,6	12,7	8,4	5,7	8,6	10,1	8
1992	13	57,3	16,6						
1993	14	61	16,1						
1994	15	64,9	15,9						
1995	16	72,4	19,9	15,9	12,2	9,9	8,6	10,3	5,7
1996	17	80	23,3	19,3	15,6	15,6	12,1		
	18	84,8	23,6	19,6	15,8	13,6	12,1		
1998	19	87,5	22,1	17,9	13,9	11,7	10,2	8,1	
1999	20	89,7	20,2	15,8	11,7	9,3	7,8	5,6	3,1
2000	21	95,2	21	16,9	12,8	10,3	8,8	6,6	4,1
001	22	100,3	21,8	17,4	13,4	10,9	9,4	7,1	4,6
متوسط الخطأ لتنبؤات $\bar{\epsilon}_{pr}$			19,4	15,9	12,4	10,6	9,9	8	5,1
عدد حدود السلسلة المدرسة $n_L$			11	12	13	14	15	16	17
عدد حدود التنبأ $L$			11	10	9	8	7	6	5

نلاحظ أن السلسلة الزمنية قسمت إلى قسمين متساوين، لكل قسم 11 حداً، ثم قمنا بتشكيل معادلة الاتجاه للقسم الثاني والتي هي:  $Y^* = 3.74 + 3.40t$ . بتعويض قيم  $t$  في هذه المعادلة من عام 1961 إلى 1971، نكون قد تنبأنا بقيم الإنتاج خلال هذه الفترة. قمنا بحساب الأخطاء المرتكبة التي دونت في خانة الحالة الأولى وفق المعادلة (6.18)، ومن ثم قمنا بحساب متوسط الأخطاء الذي يساوي:  $\bar{\epsilon}_{pr} = 19.4$ . هذه القيمة تقابل 11 حداً من حدود السلسلة المدروسة  $n_L$  و 11 حداً آخر من حدود التنبأ بها  $L$ .

نظيف حد إلى الحدود الأحدي عشرة الأولى ومن ثم تبحت عن معادلة التمثيل لـ 12 حداً ونكرر الخطوات السابقة. دونت النتائج في الجدول رقم 2 السابق. نتحصل في الأخير على الجدول التالي الذي يعطي قيم  $\bar{\epsilon}_{pr}$  بدلالة  $n_L$  و  $L$ :

$\bar{\epsilon}_{pr}$	$n_L$	$L$
19.4	11	11
15.9	12	10
12.4	13	9
10.6	14	8
9.9	15	7
8	16	6
5.1	17	5

يمكننا تشكيل معادلة خطية ثنائية (إرتباط ثنائي) إنطلاقاً من هذه البيانات فنجد:

$$\bar{\epsilon} = 3.86 - 0.44n_L + 1.74L$$

حيث أن  $R=0.98$  (معامل الارتباط المتعدد).

إن هذه المعادلة تعني أنه إذا إزداد عدد الحدود المدروسة  $n_L$  بمقدار حد واحد، فإن متوسط الخطأ  $\bar{\epsilon}$  ينقص بمقدار 0.44 % وأنه إذا إزداد مدى التنبؤ (عدد الحدود المتنبأ بها  $L$ ) بمقدار سنة واحدة فإن  $\bar{\epsilon}$  يزداد بمقدار 1.74 %.

فعندما نريد حساب كمية إنتاج الحليب لسنة 2005، فإن هذا يعني  
أخذ  $t = 25$ ، يكون:

$$Y^* = -2.85 + 4.53(25) = 110.4 \text{ (م.طن)}$$

أما مجال ثقة هذا التنبؤ هو:

$[\tilde{Y} - ZS_{yy}^*, \tilde{Y} + ZS_{yy}^*]$  وحيث أن الخطأ المرتكب في هذه الحالة غير ثابت،  
بل هو مرتبط بعدد الحدود المدروسة  $n_L$  وعدد الحدود المتنبأ بها  $L$ ، فإنه من  
الأفضل استخدام معادلة تمثيل متوسط الخطأ، أي أن:

$$\left[ Y^*_{n+L} - Z \frac{\bar{\epsilon}_{pr} Y^*}{100}, Y^*_{n+L} + Z \frac{\bar{\epsilon}_{pr} Y^*}{100} \right] \quad (6.19)$$

والذي يساوي في مثالنا (على إعتبار أن احتمال مجال الثقة هو 0.95، أي  
أن  $Z = 2$ ):

$$\left[ Y^*_{n+L} - 2x \frac{(3.68 - 0.44n_L + 1.74L)Y^*}{100}, Y^*_{n+L} + 2x \frac{(3.68 - 0.44n_L + 1.74L)Y^*}{100} \right]$$

وحيث أن  $n_L = 22$  و  $L = 5$  فإن مجال الثقة يكون:

$$\left[ 110.4 - 2x \frac{((3.68 - 0.44(22) + 1.74(5))110.4}{100}, 110.4 + 2x \frac{((3.68 - 0.44(22) + 1.74(5))110.4}{100} \right]$$

وأخيرا نجد:

$$[77.94, 142.58] \text{ باحتمال قدره: } 0.95$$

## تمارين مختارة

### التمرين الأول:

يمثل الجدول التالي، كمية اللحم المستهلكة للفرد الواحد في بلد ما.

السنة	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
كمية اللحم (كغ/فرد)	10	12	12	13	14	16	17	18	20	22

### المطلوب:

- أرسم المنحنى التاريخي لهذه السلسلة،
- أوجد معادلة الاتجاه مع حساب معامل التحديد،
- أوجد معادلة التغيرات الدورية ثم مثلها بيانيا،
- ماهي الأخطاء الناجمة عن هذا التمثيل.

### التمرين الثاني:

يمثل الجدول التالي إنتاج الجزائر من القمح، خلال 10 سنوات.

السنة	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
كمية القمح (قنطار)	1181	1035	573	702	707	926	768	632	492	585

### المطلوب:

- شكل معادلة الاتجاه مع حساب معامل التحديد،
- بعد حساب الأخطاء التي تنجم عن عملية تنبأ (بطريقة الخطأ المتزايد)، كم سيكون إنتاج القمح في سنة 2000. ماهو مجال الثقة اللازم لهذا التنبؤ، عند احتمال 0.95.



### التمرين الثالث:

يمثل الجدول كميات الغاز المستهلكة (م<sup>3</sup>) لمنطقة ما، بدلالة أربعة فصول:

الفصول السنوات	I	II	III	IV
2000	1520	2514	3251	2458
2001	1879	2697	4523	1952
2002	2045	3258	3458	1523
2003	2486	3458	4856	2698

### المطلوب:

- شكل المنحنى التاريخي لهذه السلسلة، قبل وبعد حساب المتوسطات المتحركة بطول 3، ماذا تستنتج؟.
- قدر التغيرات الموسمية لهذه السلسلة،
- قدر التغيرات العشوائية،
- شكل معادلة الاتجاه مع حساب معامل التحديد،
- كم ستكون الكمية المستهلكة من الغاز تحت نفس الظروف في سنة 2006، أحسب كل الأخطاء الناجمة عن هذا التنبؤ.



## المراجع

أ) باللغة العربية:

- ابراهيم محمد العلي (1980): مدخل في نظرية العينات. مديرية الكتب والمطبوعات الجامعية. منشورات جامعة حلب، كلية الإقتصاد والتجارة. 50 ص - 96 ص.
- ابراهيم محمد العلي (1982): نظرية الارتباط. مديرية الكتب والمطبوعات الجامعية. منشورات جامعة حلب، كلية الإقتصاد والتجارة. 250 صفحة.
- بوعبد الله صالح (2006): محاضرات الإحصاء الرياضي، لطلبة كلية العلوم الاقتصادية.
- السعدي رجال (1995): نظرية الإحتمال، مسائل وتمارين محلولة. ديوان المطبوعات الجامعية. الجزء الأول. 355 صفحة.
- موارد شبيجل (1998): الإحصاء. سلسلة ملخصات شوم، الطبعة الخامسة. الدار الدولية للنشر والتوزيع. القاهرة. مصر. 564 صفحة.
- سيمور ليبشتر (1984): الإحتمالات. سلسلة ملخصات شوم. دار اكجروهيل للنشر، دار المريخ بالملكة العربية السعودية. 210 صفحة.
- كنجو. أنيس (1979): الإحصاء الرياضي. مطبعة زيد ابن ثابت. 544 صفحة.

## ب) باللغة الأجنبية:

- CHAMBADAL.L (1978): Calcul des probabilités. Edition DUNOD.
- DAGNELIE.P (1982): Théorie et méthode statistique. Vol 1 et 2.
- GENET.J, PUPION.G et REPUSSARD.M (1974): Probabilités, statistiques et sondages. Edition Vuibert. 319 P.
- HAMDANI.H (2001): Statistique descriptive. Office des Publications Universitaires.259P.
- LEBART.L, MORINEAU.A , JABAR.N (1978): Techniques de la description statistiques, méthodes et logiciels pour l'analyse des grands tableau. Edition Dunod.351 P.
- MERCIER.M (1996): Biostatistique et probabilités. Collection PCEM. Edition Ellipses. 191 P.
- MURRAR.R.S (1981): Probabilités et statistiques. Séries Schaum. Edition McGraw – Hill. France et Canada. 385 P.
- SAPORTA.G (1982): Théories et méthodes de la statistique. Publications de l'institut français du pétrole. Société des éditions TECHNIP.386 P.

ملحق الجداول الإحصائية:

جدول 1: احتمالات ذو الحدين:  $P(X = x) = C_n^x p^x q^{n-x} = \frac{n!}{(n-x)!x!} p^x q^{n-x}$

n = 10											
p		0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
k	0	0,5987	0,3487	0,1969	0,1074	0,0563	0,0282	0,0135	0,0060	0,0025	0,0010
	1	0,9139	0,7361	0,5443	0,3758	0,2440	0,1493	0,0860	0,0464	0,0233	0,0107
	2	0,9885	0,9298	0,8202	0,6778	0,5256	0,3828	0,2616	0,1673	0,0996	0,0547
	3	0,9990	0,9872	0,9500	0,8791	0,7759	0,6496	0,5138	0,3823	0,2660	0,1719
	4	0,9999	0,9984	0,9901	0,9672	0,9219	0,8497	0,7515	0,6331	0,5044	0,3770
	5	1,0000	0,9999	0,9986	0,9936	0,9803	0,9527	0,9051	0,8338	0,7384	0,6230
	6	1,0000	1,0000	0,9999	0,9991	0,9965	0,9894	0,9740	0,9432	0,8980	0,8281
	7	1,0000	1,0000	1,0000	0,9999	0,9996	0,9984	0,9952	0,9877	0,9726	0,9453
	8	1,0000	1,0000	1,0000	1,0000	1,0000	0,9999	0,9995	0,9983	0,9955	0,9893
	9	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9999	0,9997	0,9990
	10	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

n=20											
p		0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
k	0	0,3585	0,1216	0,0388	0,0115	0,0032	0,0008	0,0002	0,0000	0,0000	0,0000
	1	0,7358	0,3917	0,1756	0,0692	0,0243	0,0076	0,0021	0,0005	0,0001	0,0000
	2	0,9245	0,6769	0,4049	0,2061	0,0913	0,0355	0,0121	0,0036	0,0009	0,0002
	3	0,9841	0,8670	0,6477	0,4114	0,2252	0,1071	0,0444	0,0160	0,0049	0,0013
	4	0,9974	0,9568	0,8298	0,6296	0,4148	0,2375	0,1182	0,0510	0,0189	0,0059
	5	0,9997	0,9887	0,9327	0,8042	0,6172	0,4164	0,2454	0,1256	0,0553	0,0207
	6	1,0000	0,9976	0,9781	0,9133	0,7858	0,6080	0,4166	0,2500	0,1299	0,0577
	7	1,0000	0,9996	0,9941	0,9679	0,8982	0,7723	0,6010	0,4159	0,2520	0,1316
	8	1,0000	0,9999	0,9987	0,9900	0,9591	0,8867	0,7624	0,5956	0,4143	0,2517
	9	1,0000	1,0000	0,9998	0,9974	0,9861	0,9520	0,8782	0,7553	0,5914	0,4119
	10	1,0000	1,0000	1,0000	0,9994	0,9961	0,9829	0,9468	0,8725	0,7507	0,5881
	11	1,0000	1,0000	1,0000	0,9999	0,9991	0,9949	0,9804	0,9435	0,8692	0,7483
	12	1,0000	1,0000	1,0000	1,0000	0,9998	0,9987	0,9940	0,9790	0,9420	0,8684
	13	1,0000	1,0000	1,0000	1,0000	1,0000	0,9997	0,9985	0,9935	0,9786	0,9423
	14	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9997	0,9984	0,9936	0,9793
	15	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9997	0,9983	0,9941
	16	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9997	0,9987
	17	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9998
	18	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

جدول 2 - احتمالات بواسون  $P(X = k) = \frac{m^k e^{-m}}{k!}$

k	Probabilités individuelles $Pr(k) = e^{-m} \frac{m^k}{k!}$								
	m = 1,0	m = 1,5	m = 2,0	m = 2,5	m = 3,0	m = 3,5	m = 4,0	m = 4,5	m = 5,0
0	0,3679	0,2231	0,1353	0,0821	0,0498	0,0302	0,0183	0,0111	0,0067
1	0,3679	0,3347	0,2707	0,2052	0,1494	0,1057	0,0733	0,0500	0,0337
2	0,1839	0,2510	0,2707	0,2565	0,2240	0,1850	0,1465	0,1125	0,0842
3	0,0613	0,1253	0,1804	0,2138	0,2240	0,2158	0,1934	0,1687	0,1404
4	0,0153	0,0471	0,0902	0,1336	0,1680	0,1888	0,1954	0,1898	0,1755
5	0,0031	0,0141	0,0361	0,0668	0,1008	0,1322	0,1563	0,1708	0,1755
6	0,0005	0,0035	0,0120	0,0278	0,0504	0,0771	0,1042	0,1281	0,1462
7	0,0001	0,0008	0,0034	0,0099	0,0216	0,0385	0,0595	0,0824	0,1044
8		0,0001	0,0009	0,0031	0,0081	0,0169	0,0298	0,0483	0,0653
9			0,0002	0,0009	0,0027	0,0066	0,0132	0,0232	0,0363
10				0,0002	0,0008	0,0023	0,0053	0,0104	0,0181
11					0,0002	0,0007	0,0019	0,0043	0,0082
12					0,0001	0,0002	0,0006	0,0016	0,0034
13						0,0001	0,0002	0,0006	0,0013
14							0,0001	0,0002	0,0005
15								0,0001	0,0002
16									0,0001

k	Probabilités individuelles $Pr(k) = e^{-m} \frac{m^k}{k!}$								
	m = 5,5	m = 6,0	m = 6,5	m = 7,0	m = 7,5	m = 8,0	m = 8,5	m = 9,0	m = 9,5
0	0,0041	0,0025	0,0015	0,0009	0,0006	0,0003	0,0002	0,0001	0,0001
1	0,0223	0,0148	0,0098	0,0064	0,0041	0,0027	0,0017	0,0011	0,0007
2	0,0618	0,0446	0,0318	0,0223	0,0156	0,0107	0,0074	0,0050	0,0034
3	0,1133	0,0892	0,0688	0,0521	0,0389	0,0286	0,0208	0,0150	0,0107
4	0,1558	0,1338	0,1118	0,0912	0,0728	0,0573	0,0443	0,0337	0,0254
5	0,1714	0,1606	0,1454	0,1277	0,1094	0,0916	0,0752	0,0607	0,0483
6	0,1571	0,1606	0,1575	0,1490	0,1367	0,1221	0,1066	0,0911	0,0764
7	0,1234	0,1377	0,1462	0,1490	0,1465	0,1386	0,1294	0,1171	0,1037
8	0,0849	0,1033	0,1188	0,1304	0,1373	0,1396	0,1375	0,1318	0,1232
9	0,0519	0,0688	0,0838	0,1014	0,1144	0,1241	0,1299	0,1318	0,1300
10	0,0285	0,0413	0,0558	0,0710	0,0858	0,0993	0,1104	0,1186	0,1235
11	0,0143	0,0225	0,0330	0,0452	0,0585	0,0722	0,0853	0,0970	0,1067
12	0,0065	0,0113	0,0179	0,0264	0,0366	0,0481	0,0604	0,0728	0,0844
13	0,0028	0,0052	0,0089	0,0142	0,0211	0,0296	0,0393	0,0504	0,0617
14	0,0011	0,0022	0,0041	0,0071	0,0113	0,0169	0,0240	0,0324	0,0419
15	0,0004	0,0009	0,0018	0,0033	0,0057	0,0090	0,0138	0,0194	0,0255
16	0,0001	0,0003	0,0007	0,0014	0,0026	0,0043	0,0072	0,0108	0,0157
17		0,0001	0,0003	0,0006	0,0012	0,0021	0,0036	0,0058	0,0088
18			0,0001	0,0002	0,0005	0,0009	0,0017	0,0029	0,0046
19				0,0001	0,0002	0,0004	0,0008	0,0014	0,0023
20					0,0001	0,0002	0,0003	0,0006	0,0011
21						0,0001	0,0001	0,0003	0,0005
22							0,0001	0,0001	0,0002
23								0,0001	0,0001
24									0,0001

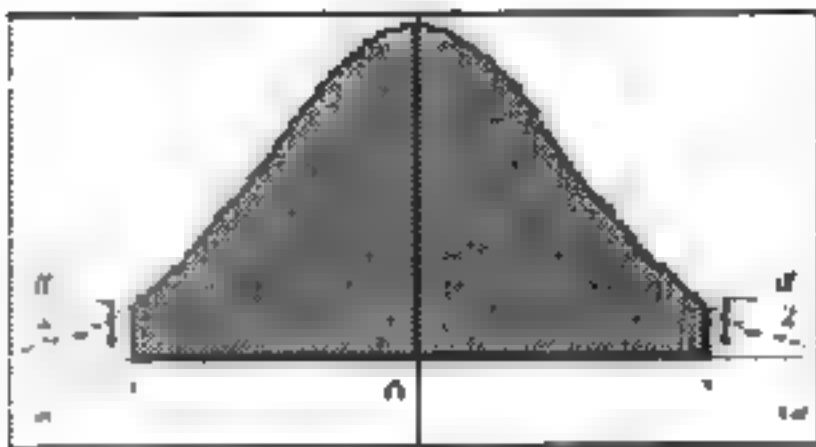




جدول 3 - احتمالات التوزيع الطبيعي القياسي.

$Z$	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1916	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.267	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3926	.3944	.3962	.3980	.3997	.4016
1.3	.4032	.4049	.4068	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4506	.4515	.4525	.4535	.4545
1.7	.4564	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

جدول 4 - احتمالات توزيع t.



$\alpha$	1	0.8	0.6	0.4	0.2	0.1	0.05	0.02	0.01	0.002	0.001
$1 - \alpha$	0	0.2	0.4	0.6	0.8	0.9	0.95	0.98	0.99	0.998	0.999
$\nu = dd$											
1	0.0000	0.3249	0.7265	0.9370	0.9777	0.9933	0.9975	0.9993	0.9999	1.0000	1.0000
2	0.0000	0.2887	0.6872	0.9307	0.9750	0.9900	0.9950	0.9975	0.9990	0.9995	0.9997
3	0.0000	0.2767	0.6844	0.9283	0.9737	0.9893	0.9943	0.9968	0.9983	0.9990	0.9993
4	0.0000	0.2707	0.6806	0.9259	0.9722	0.9880	0.9930	0.9955	0.9970	0.9980	0.9985
5	0.0000	0.2672	0.6784	0.9243	0.9715	0.9873	0.9923	0.9948	0.9963	0.9975	0.9980
6	0.0000	0.2648	0.6765	0.9231	0.9708	0.9865	0.9915	0.9940	0.9955	0.9968	0.9973
7	0.0000	0.2632	0.6751	0.9221	0.9702	0.9858	0.9908	0.9933	0.9948	0.9961	0.9966
8	0.0000	0.2619	0.6740	0.9212	0.9696	0.9852	0.9902	0.9927	0.9942	0.9955	0.9960
9	0.0000	0.2610	0.6733	0.9205	0.9690	0.9846	0.9896	0.9921	0.9936	0.9949	0.9954
10	0.0000	0.2602	0.6726	0.9199	0.9684	0.9840	0.9890	0.9915	0.9930	0.9943	0.9948
11	0.0000	0.2596	0.6720	0.9193	0.9678	0.9834	0.9884	0.9909	0.9924	0.9937	0.9942
12	0.0000	0.2590	0.6714	0.9187	0.9672	0.9828	0.9878	0.9903	0.9918	0.9931	0.9936
13	0.0000	0.2586	0.6709	0.9182	0.9666	0.9822	0.9872	0.9900	0.9915	0.9928	0.9933
14	0.0000	0.2582	0.6704	0.9177	0.9660	0.9816	0.9866	0.9896	0.9911	0.9924	0.9929
15	0.0000	0.2579	0.6700	0.9172	0.9655	0.9810	0.9860	0.9890	0.9905	0.9918	0.9923
16	0.0000	0.2576	0.6696	0.9167	0.9650	0.9804	0.9854	0.9884	0.9899	0.9912	0.9917
17	0.0000	0.2573	0.6692	0.9162	0.9645	0.9798	0.9848	0.9878	0.9893	0.9906	0.9911
18	0.0000	0.2571	0.6688	0.9158	0.9640	0.9792	0.9842	0.9872	0.9887	0.9900	0.9905
19	0.0000	0.2569	0.6684	0.9153	0.9635	0.9786	0.9836	0.9866	0.9881	0.9894	0.9899
20	0.0000	0.2567	0.6680	0.9149	0.9630	0.9780	0.9830	0.9860	0.9875	0.9888	0.9893
21	0.0000	0.2566	0.6676	0.9144	0.9625	0.9774	0.9824	0.9854	0.9869	0.9882	0.9887
22	0.0000	0.2564	0.6672	0.9140	0.9620	0.9768	0.9818	0.9848	0.9863	0.9876	0.9881
23	0.0000	0.2563	0.6668	0.9135	0.9615	0.9762	0.9812	0.9842	0.9857	0.9870	0.9875
24	0.0000	0.2562	0.6664	0.9131	0.9610	0.9756	0.9806	0.9836	0.9851	0.9864	0.9869
25	0.0000	0.2561	0.6660	0.9126	0.9605	0.9750	0.9800	0.9830	0.9845	0.9858	0.9863
26	0.0000	0.2560	0.6656	0.9122	0.9600	0.9744	0.9794	0.9824	0.9839	0.9852	0.9857
27	0.0000	0.2559	0.6652	0.9117	0.9595	0.9738	0.9788	0.9818	0.9833	0.9846	0.9851
28	0.0000	0.2558	0.6648	0.9113	0.9590	0.9732	0.9782	0.9812	0.9827	0.9840	0.9845
29	0.0000	0.2557	0.6644	0.9108	0.9585	0.9726	0.9776	0.9806	0.9821	0.9834	0.9839
30	0.0000	0.2556	0.6640	0.9104	0.9580	0.9720	0.9770	0.9800	0.9815	0.9828	0.9833
40	0.0000	0.2550	0.6632	0.9096	0.9572	0.9712	0.9762	0.9792	0.9807	0.9820	0.9825
50	0.0000	0.2547	0.6628	0.9092	0.9568	0.9708	0.9758	0.9788	0.9803	0.9816	0.9821
60	0.0000	0.2545	0.6624	0.9088	0.9564	0.9704	0.9754	0.9784	0.9799	0.9812	0.9817
70	0.0000	0.2543	0.6620	0.9084	0.9560	0.9700	0.9750	0.9780	0.9795	0.9808	0.9813
80	0.0000	0.2542	0.6616	0.9080	0.9556	0.9696	0.9746	0.9776	0.9791	0.9804	0.9809
90	0.0000	0.2541	0.6613	0.9076	0.9552	0.9692	0.9742	0.9772	0.9787	0.9800	0.9805
100	0.0000	0.2540	0.6610	0.9072	0.9548	0.9688	0.9738	0.9768	0.9783	0.9796	0.9801
200	0.0000	0.2537	0.6605	0.9067	0.9543	0.9683	0.9733	0.9763	0.9778	0.9791	0.9796
$\infty$	0.0000	0.2533	0.6600	0.9063	0.9539	0.9679	0.9729	0.9759	0.9774	0.9787	0.9792

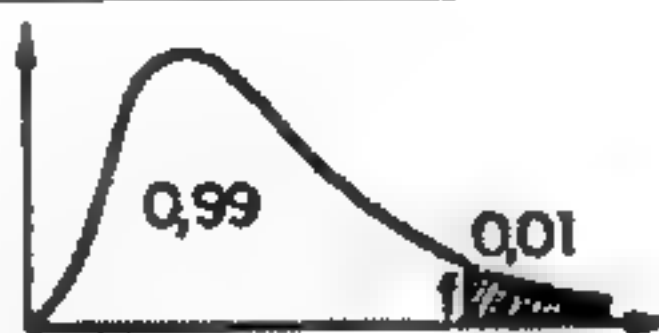


جدول 5 - احتمالات توزيع F عند  $\alpha = 5\%$ .

$\nu_1 \backslash \nu_2$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	161	200	216	225	230	234	237	239	241	242	243	244	245	245	246	246	247	247
2	18.5	19.0	19.2	19.3	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70	8.69	8.68	8.67
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86	5.84	5.83	5.82
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62	4.60	4.59	4.58
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94	3.92	3.91	3.90
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51	3.49	3.48	3.47
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22	3.20	3.19	3.17
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96
10	4.90	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.68	2.67
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.48
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46	2.44	2.43	2.41
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40	2.38	2.37	2.35
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35	2.33	2.32	2.30
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31	2.29	2.27	2.26
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27	2.25	2.23	2.22
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.21	2.20	2.18
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.22	2.20	2.18	2.17	2.15
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.22	2.20	2.18	2.16	2.14	2.12
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.20	2.17	2.15	2.13	2.11	2.10
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.23	2.20	2.18	2.15	2.13	2.11	2.09	2.07
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.21	2.18	2.15	2.13	2.11	2.09	2.07	2.05
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.20	2.16	2.14	2.11	2.09	2.07	2.05	2.04
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.12	2.09	2.07	2.05	2.03	2.02
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.17	2.13	2.10	2.08	2.06	2.04	2.02	2.00
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.09	2.06	2.04	2.02	2.00	1.99
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.08	2.05	2.03	2.01	1.99	1.97
30	4.17	3.32	2.92	2.69	2.54	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.06	2.04	2.01	1.99	1.98	1.96
32	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.19	2.14	2.10	2.07	2.04	2.01	1.99	1.97	1.95	1.94
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17	2.12	2.08	2.05	2.02	1.99	1.97	1.95	1.93	1.92
36	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.15	2.11	2.07	2.03	2.00	1.98	1.95	1.93	1.92	1.90
38	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.99	1.96	1.94	1.92	1.90	1.88
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.97	1.95	1.92	1.90	1.89	1.87
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.03	1.99	1.96	1.93	1.91	1.89	1.87	1.86
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.98	1.95	1.92	1.90	1.88	1.86	1.84
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.15	2.09	2.04	2.00	1.97	1.94	1.91	1.89	1.87	1.85	1.83
48	4.04	3.19	2.80	2.56	2.41	2.29	2.21	2.14	2.08	2.03	1.99	1.96	1.93	1.90	1.88	1.86	1.84	1.82
50	4.03	3.18	2.79	2.55	2.40	2.28	2.20	2.13	2.07	2.03	1.99	1.95	1.92	1.89	1.87	1.85	1.83	1.81
55	4.02	3.16	2.77	2.54	2.38	2.27	2.18	2.11	2.06	2.01	1.97	1.93	1.90	1.88	1.85	1.83	1.81	1.79
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.89	1.86	1.84	1.82	1.80	1.78
65	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.03	1.98	1.94	1.90	1.87	1.85	1.82	1.80	1.78	1.76
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.02	1.97	1.93	1.89	1.86	1.84	1.81	1.79	1.77	1.75
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00	1.95	1.91	1.88	1.84	1.82	1.79	1.77	1.75	1.73
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94	1.90	1.86	1.83	1.80	1.78	1.76	1.74	1.72
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.89	1.85	1.82	1.79	1.77	1.75	1.73	1.71
125	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.96	1.91	1.87	1.83	1.80	1.77	1.75	1.72	1.70	1.69
150	3.90	3.06	2.66	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.85	1.82	1.79	1.76	1.73	1.71	1.69	1.67
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.84	1.80	1.77	1.74	1.72	1.69	1.67	1.66
300	3.87	3.03	2.63	2.40	2.24	2.13	2.04	1.97	1.91	1.86	1.82	1.78	1.75	1.72	1.70	1.68	1.66	1.64
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.77	1.74	1.71	1.69	1.66	1.64	1.62
1000	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89	1.84	1.80	1.76	1.73	1.70	1.68	1.65	1.63	1.61
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.79	1.75	1.72	1.69	1.67	1.64	1.62	1.60

# جدول 6 - احتمالات توزيع F عند $\alpha = 1\%$ .

مثلا: عند درجة حرية  $v_1=9, v_2=11$  فإن قيمة  $F = 4.63$



$v_2 \backslash v_1$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	(Les valeurs de la première ligne doivent être multipliées par 10)																	
1	405	300	240	200	170	150	135	125	115	105	100	95	90	85	80	75	70	65
2	98.5	99.0	99.2	99.3	99.4	99.5	99.6	99.7	99.8	99.9	100.0	100.1	100.2	100.3	100.4	100.5	100.6	100.7
3	33.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	27.0	26.9	26.9	26.9	26.8	26.8	26.8
4	22.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.4	14.3	14.2	14.2	14.2	14.1	14.1
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.96	9.89	9.82	9.77	9.72	9.68	9.64	9.61
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72	7.66	7.60	7.56	7.52	7.48	7.45
7	12.2	9.35	8.13	7.50	7.10	6.82	6.61	6.44	6.32	6.22	6.14	6.07	6.01	5.96	5.92	5.88	5.84	5.81
8	11.3	8.65	7.39	6.76	6.36	6.08	5.87	5.70	5.58	5.48	5.40	5.33	5.27	5.22	5.18	5.14	5.10	5.07
9	10.6	8.02	6.76	6.13	5.73	5.45	5.24	5.07	4.95	4.85	4.77	4.70	4.64	4.60	4.56	4.52	4.48	4.46
10	10.0	7.36	6.10	5.47	5.07	4.79	4.58	4.41	4.29	4.19	4.11	4.04	3.98	3.94	3.90	3.86	3.82	3.80
11	9.65	7.21	5.95	5.32	4.92	4.64	4.43	4.26	4.14	4.04	3.96	3.89	3.83	3.79	3.75	3.71	3.67	3.65
12	9.33	6.93	5.67	5.04	4.64	4.36	4.15	3.98	3.86	3.76	3.68	3.61	3.55	3.51	3.47	3.43	3.39	3.37
13	9.17	6.70	5.44	4.81	4.41	4.13	3.92	3.75	3.63	3.53	3.45	3.38	3.32	3.28	3.24	3.20	3.16	3.14
14	8.96	6.51	5.25	4.62	4.22	3.94	3.73	3.56	3.44	3.34	3.26	3.19	3.13	3.09	3.05	3.01	2.97	2.95
15	8.78	6.36	5.10	4.47	4.07	3.79	3.58	3.41	3.29	3.19	3.11	3.04	2.98	2.94	2.90	2.86	2.82	2.80
16	8.63	6.23	4.97	4.34	3.94	3.66	3.45	3.28	3.16	3.06	2.98	2.91	2.85	2.81	2.77	2.73	2.69	2.67
17	8.49	6.11	4.85	4.22	3.82	3.54	3.33	3.16	3.04	2.94	2.86	2.79	2.73	2.69	2.65	2.61	2.57	2.55
18	8.37	6.01	4.75	4.12	3.72	3.44	3.23	3.06	2.94	2.84	2.76	2.69	2.63	2.59	2.55	2.51	2.47	2.45
19	8.26	5.93	4.67	4.04	3.64	3.36	3.15	2.98	2.86	2.76	2.68	2.61	2.55	2.51	2.47	2.43	2.39	2.37
20	8.16	5.85	4.59	3.96	3.56	3.28	3.07	2.90	2.78	2.68	2.60	2.53	2.47	2.43	2.39	2.35	2.31	2.29
21	8.07	5.78	4.52	3.89	3.49	3.21	3.00	2.83	2.71	2.61	2.53	2.46	2.40	2.36	2.32	2.28	2.24	2.22
22	7.99	5.72	4.46	3.83	3.43	3.15	2.94	2.77	2.65	2.55	2.47	2.40	2.34	2.30	2.26	2.22	2.18	2.16
23	7.92	5.66	4.40	3.77	3.37	3.09	2.88	2.71	2.59	2.49	2.41	2.34	2.28	2.24	2.20	2.16	2.12	2.10
24	7.86	5.61	4.35	3.72	3.32	3.04	2.83	2.66	2.54	2.44	2.36	2.29	2.23	2.19	2.15	2.11	2.07	2.05
25	7.77	5.57	4.31	3.68	3.28	3.00	2.79	2.62	2.50	2.40	2.32	2.25	2.19	2.15	2.11	2.07	2.03	2.01
26	7.72	5.53	4.27	3.64	3.24	2.96	2.75	2.58	2.46	2.36	2.28	2.21	2.15	2.11	2.07	2.03	1.99	1.97
27	7.68	5.49	4.23	3.60	3.20	2.92	2.71	2.54	2.42	2.32	2.24	2.17	2.11	2.07	2.03	1.99	1.95	1.93
28	7.64	5.45	4.19	3.56	3.16	2.88	2.67	2.50	2.38	2.28	2.20	2.13	2.07	2.03	1.99	1.95	1.91	1.89
29	7.60	5.42	4.16	3.53	3.13	2.85	2.64	2.47	2.35	2.25	2.17	2.10	2.04	2.00	1.96	1.92	1.88	1.86
30	7.56	5.39	4.13	3.50	3.10	2.82	2.61	2.44	2.32	2.22	2.14	2.07	2.01	1.97	1.93	1.89	1.85	1.83
32	7.50	5.34	4.08	3.45	3.05	2.77	2.56	2.39	2.27	2.17	2.09	2.02	1.96	1.92	1.88	1.84	1.80	1.78
34	7.44	5.29	4.03	3.40	3.00	2.72	2.51	2.34	2.22	2.12	2.04	1.97	1.91	1.87	1.83	1.79	1.75	1.73
36	7.40	5.25	4.00	3.37	2.97	2.69	2.48	2.31	2.19	2.09	2.01	1.94	1.88	1.84	1.80	1.76	1.72	1.70
38	7.35	5.21	3.95	3.32	2.92	2.64	2.43	2.26	2.14	2.04	1.96	1.89	1.83	1.79	1.75	1.71	1.67	1.65
40	7.31	5.18	3.92	3.29	2.89	2.61	2.40	2.23	2.11	2.01	1.93	1.86	1.80	1.76	1.72	1.68	1.64	1.62
42	7.28	5.15	3.89	3.26	2.86	2.58	2.37	2.20	2.08	1.98	1.90	1.83	1.77	1.73	1.69	1.65	1.61	1.59
44	7.25	5.12	3.86	3.23	2.83	2.55	2.34	2.17	2.05	1.95	1.87	1.80	1.74	1.70	1.66	1.62	1.58	1.56
46	7.22	5.10	3.84	3.21	2.81	2.53	2.32	2.15	2.03	1.93	1.85	1.78	1.72	1.68	1.64	1.60	1.56	1.54
48	7.19	5.08	3.82	3.19	2.79	2.51	2.30	2.13	2.01	1.91	1.83	1.76	1.70	1.66	1.62	1.58	1.54	1.52
50	7.17	5.06	3.80	3.17	2.77	2.49	2.28	2.11	1.99	1.89	1.81	1.74	1.68	1.64	1.60	1.56	1.52	1.50
55	7.12	5.01	3.76	3.13	2.73	2.45	2.24	2.07	1.95	1.85	1.77	1.70	1.64	1.60	1.56	1.52	1.48	1.46
60	7.08	4.98	3.72	3.09	2.69	2.41	2.20	2.03	1.91	1.81	1.73	1.66	1.60	1.56	1.52	1.48	1.44	1.42
65	7.04	4.95	3.68	3.05	2.65	2.37	2.16	1.99	1.87	1.77	1.69	1.62	1.56	1.52	1.48	1.44	1.40	1.38
70	7.01	4.92	3.65	3.02	2.62	2.34	2.13	1.96	1.84	1.74	1.66	1.59	1.53	1.49	1.45	1.41	1.37	1.35
80	6.96	4.88	3.60	2.97	2.57	2.29	2.08	1.91	1.79	1.69	1.61	1.54	1.48	1.44	1.40	1.36	1.32	1.30
90	6.93	4.85	3.57	2.94	2.54	2.26	2.05	1.88	1.76	1.66	1.58	1.51	1.45	1.41	1.37	1.33	1.29	1.27
100	6.90	4.82	3.54	2.91	2.51	2.23	2.02	1.85	1.73	1.63	1.55	1.48	1.42	1.38	1.34	1.30	1.26	1.24
125	6.84	4.78	3.49	2.86	2.46	2.18	1.97	1.80	1.68	1.58	1.50	1.43	1.37	1.33	1.29	1.25	1.21	1.19
150	6.81	4.75	3.46	2.83	2.43	2.15	1.94	1.77	1.65	1.55	1.47	1.40	1.34	1.30	1.26	1.22	1.18	1.16
200	6.76	4.71	3.41	2.78	2.38	2.10	1.89	1.72	1.60	1.50	1.42	1.35	1.29	1.25	1.21	1.17	1.13	1.11
300	6.72	4.68	3.38	2.75	2.35	2.07	1.86	1.69	1.57	1.47	1.39	1.32	1.26	1.22	1.18	1.14	1.10	1.08
500	6.69	4.65	3.35	2.72	2.32	2.04	1.83	1.66	1.54	1.44	1.36	1.29	1.23	1.19	1.15	1.11	1.07	1.05
1000	6.66	4.63	3.33	2.70	2.30	2.02	1.81	1.64	1.52	1.42	1.34	1.27	1.21	1.17	1.13	1.09	1.05	1.03
∞	6.63	4.61	3.32	2.69	2.29	2.01	1.80	1.63	1.51	1.41	1.33	1.26	1.20	1.16	1.12	1.08	1.04	1.02

## جدول 7 - احتمالات توزيع $\chi^2$

مثلا: عند درجة حرية 11، ومساحة  $p = 0.90$  أي  $\alpha = 0.10$  فإن

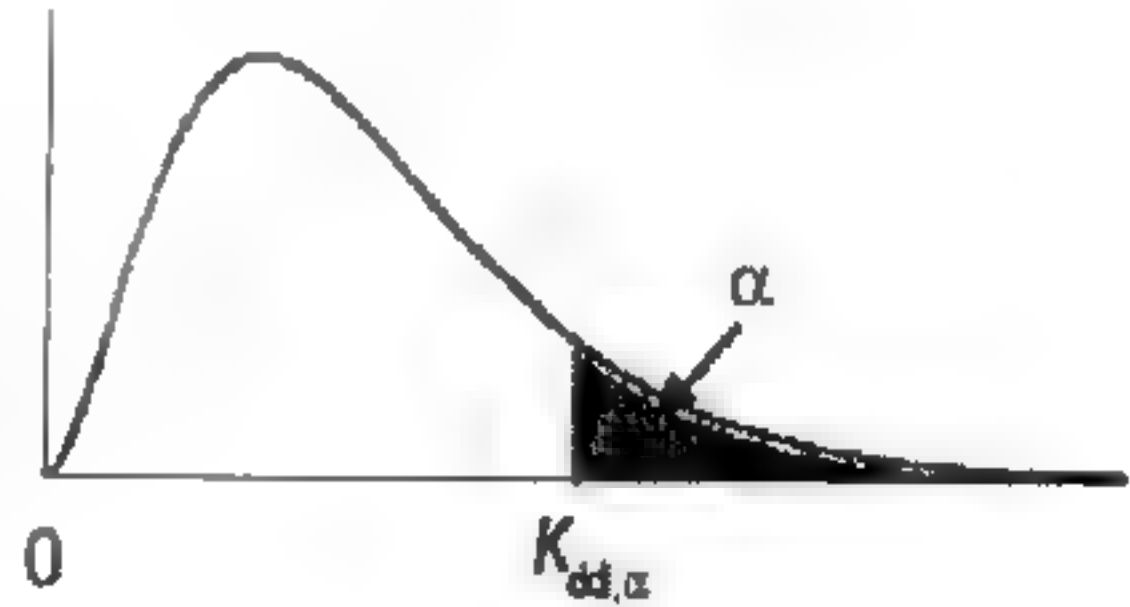
$$\chi^2 = 17.275$$

La table donne la probabilité  $\alpha$  pour que  $\chi^2$  égale ou dépasse une valeur donnée, en fonction du nombre de degrés de liberté (d. d. l.)

Quand le nombre de degrés de liberté est élevé,

$\sqrt{2\chi^2}$  est à peu près distribué normalement

autour de  $\sqrt{2(d.d.l.) - 1}$  avec une variance égale à 1



$\alpha$	0.90	0.50	0.30	0.20	0.10	0.05	0.02	0.01	0.001
ddl									
1	0.0158	0.455	1.074	1.642	2.706	3.841	5.412	6.635	10.827
2	0.211	1.386	2.408	3.219	4.605	5.991	7.824	9.210	13.815
3	0.584	2.366	3.665	4.642	6.251	7.815	9.837	11.345	16.266
4	1.064	3.357	4.878	5.989	7.779	9.488	11.668	13.277	18.467
5	1.610	4.351	6.064	7.289	9.236	11.070	13.388	15.086	20.515
6	2.204	5.348	7.231	8.558	10.645	12.592	15.033	16.812	22.457
7	2.833	6.346	8.383	9.803	12.017	14.067	16.622	18.475	24.322
8	3.490	7.344	9.524	11.030	13.362	15.507	18.168	20.090	26.125
9	4.168	8.343	10.656	12.242	14.684	16.919	19.679	21.666	27.877
10	4.865	9.342	11.781	13.442	15.987	18.307	21.161	23.209	29.588
11	5.578	10.341	12.899	14.631	17.275	19.675	22.618	24.725	31.264
12	6.304	11.340	14.011	15.812	18.549	21.026	24.054	26.217	32.909
13	7.042	12.340	15.119	16.985	19.812	22.362	25.472	27.688	34.528
14	7.790	13.339	16.222	18.151	21.064	23.685	26.873	29.141	36.123
15	8.547	14.339	17.322	19.311	22.307	24.996	28.259	30.578	37.697
16	9.312	15.338	18.418	20.465	23.542	26.296	29.633	32.000	39.252
17	10.085	16.338	19.511	21.615	24.769	27.587	30.995	33.409	40.790
18	10.865	17.338	20.601	22.760	25.989	28.869	32.346	34.805	42.312
19	11.651	18.338	21.689	23.900	27.204	30.144	33.687	36.191	43.820
20	12.443	19.337	22.775	25.038	28.412	31.410	35.020	37.566	45.315
21	13.240	20.337	23.858	26.171	29.615	32.671	36.343	38.932	46.797
22	14.041	21.337	24.939	27.301	30.813	33.924	37.659	40.289	48.268
23	14.848	22.337	26.018	28.429	32.007	35.172	38.968	41.638	49.728
24	15.659	23.337	27.096	29.553	33.196	36.415	40.270	42.980	51.179
25	16.473	24.337	28.172	30.675	34.382	37.652	41.566	44.314	52.620
26	17.292	25.336	29.246	31.795	35.563	38.885	42.856	45.642	54.052
27	18.114	26.336	30.319	32.912	36.741	40.113	44.140	46.963	55.476
28	18.939	27.336	31.391	34.027	37.916	41.337	45.419	48.278	56.893
29	19.768	28.336	32.461	35.139	39.087	42.557	46.693	49.588	58.302
30	20.599	29.336	33.530	36.250	40.256	43.773	47.962	50.892	59.703



## جدول 8 - جدول قيم Z المقابلة لقيم r.

تقرأ قيم Z من خلال السطر ثم العمود، وبداخل الجدول تقرأ قيم r

مثلا: من أجل  $r = 0.821$  نجد  $Z = 1.16$

$$r = \frac{e^{2x} - 1}{e^{2x} + 1}$$

$$z = \frac{1}{2} \log \frac{1+r}{1-r}$$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0100	0,0200	0,0300	0,0400	0,0500	0,0599	0,0699	0,0798	0,0898
0,1	0,0997	0,1098	0,1194	0,1293	0,1391	0,1489	0,1586	0,1684	0,1781	0,1877
0,2	0,1974	0,2070	0,2165	0,2260	0,2355	0,2449	0,2548	0,2638	0,2729	0,2821
0,3	0,2913	0,3004	0,3095	0,3185	0,3275	0,3364	0,3452	0,3540	0,3627	0,3714
0,4	0,3800	0,3885	0,3969	0,4053	0,4136	0,4219	0,4301	0,4382	0,4462	0,4542
0,5	0,4621	0,4699	0,4777	0,4854	0,4930	0,5005	0,5080	0,5154	0,5227	0,5299
0,6	0,5370	0,5441	0,5511	0,5580	0,5649	0,5717	0,5784	0,5850	0,5915	0,5980
0,7	0,6044	0,6107	0,6169	0,6231	0,6291	0,6351	0,6411	0,6469	0,6527	0,6584
0,8	0,6640	0,6698	0,6751	0,6805	0,6858	0,6911	0,6963	0,7014	0,7064	0,7114
0,9	0,7163	0,7211	0,7259	0,7306	0,7352	0,7397	0,7443	0,7487	0,7531	0,7574
1,0	0,7616	0,7658	0,7699	0,7739	0,7779	0,7818	0,7857	0,7895	0,7932	0,7969
1,1	0,8005	0,8041	0,8076	0,8110	0,8144	0,8178	0,8210	0,8243	0,8275	0,8306
1,2	0,8337	0,8367	0,8397	0,8426	0,8455	0,8483	0,8511	0,8538	0,8565	0,8591
1,3	0,8617	0,8643	0,8668	0,8692	0,8717	0,8741	0,8764	0,8787	0,8810	0,8832
1,4	0,8854	0,8875	0,8896	0,8917	0,8937	0,8957	0,8977	0,8996	0,9015	0,9033
1,5	0,9051	0,9069	0,9087	0,9104	0,9121	0,9138	0,9154	0,9170	0,9186	0,9201
1,6	0,9217	0,9232	0,9246	0,9261	0,9275	0,9289	0,9302	0,9316	0,9329	0,9341
1,7	0,9354	0,9368	0,9379	0,9391	0,9402	0,9414	0,9425	0,9436	0,9447	0,9458
1,8	0,9468	0,9478	0,9488	0,9498	0,9508	0,9517	0,9526	0,9535	0,9544	0,9553
1,9	0,9562	0,9570	0,9579	0,9587	0,9595	0,9603	0,9610	0,9618	0,9625	0,9631
2,0	0,9640	0,9647	0,9654	0,9660	0,9667	0,9673	0,9680	0,9686	0,9692	0,9698
2,1	0,9704	0,9710	0,9715	0,9721	0,9726	0,9732	0,9737	0,9742	0,9747	0,9752
2,2	0,9757	0,9762	0,9768	0,9773	0,9778	0,9783	0,9788	0,9792	0,9797	0,9801
2,3	0,9806	0,9810	0,9815	0,9819	0,9824	0,9828	0,9833	0,9837	0,9841	0,9845
2,4	0,9850	0,9854	0,9858	0,9862	0,9866	0,9870	0,9874	0,9878	0,9882	0,9886
2,5	0,9890	0,9894	0,9898	0,9902	0,9906	0,9910	0,9914	0,9918	0,9922	0,9926
2,6	0,9930	0,9934	0,9938	0,9942	0,9946	0,9950	0,9954	0,9958	0,9962	0,9966
2,7	0,9970	0,9974	0,9978	0,9982	0,9986	0,9990	0,9994	0,9998	0,9999	0,9999
2,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
2,9	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

مصدر الجداول الإحصائية:

SAPORTA.G (1982): Théories et méthodes de la statistique. Publications de l'institut français du pétrole. Société des éditions TECHNIP.386 P.

## الفهرس

5 .....مقدمة الطبعة الثانية:

### الفصل الأول

#### نظرية الاحتمالات

7 .....1.1-مبادئ الحساب الاحتمالي:

7 .....1.1.1- مفاهيم أساسية:

8 .....2.1.1- تعريف الاحتمال:

10 .....3.1.1-خواص الاحتمال:

10 .....4.1.1-الخواص الأساسية في نظرية الاحتمال:

10 .....(أ) - قاعدة الجمع للأحداث المتنافية:

11 .....(ب) - قاعدة الجمع للأحداث الغير المتنافية:

11 .....(ج) - قاعدة الضرب للأحداث المستقلة:

13 .....(د) - قاعدة الضرب للأحداث المرتبطة (الاحتمال الشرطي):

14 .....(هـ) - نظرية بايز:

16 .....2.1-المتغيرات العشوائية و التوزيعات الاحتمالية:

16 .....1.2.1- المتغيرات العشوائية:

17 .....1.1.2.1- أنواع المتغيرات العشوائية:

17 .....(أ) - المتغير العشوائي المنفصل:

19 .....(ب) - المتغير العشوائي المتصل:

23 .....• دالة التوزيع  $F(x)$  للمتغيرة العشوائية المتقطعة . . . . .

25 .....• دالة التوزيع  $F(x)$  للمتغير العشوائي المستمر . . . . .

27 .....3.1-التوقع الرياضي:

28 .....(أ) - حالة التوزيع الاحتمالي المنفصل:

29 .....(ب) - حالة التوزيع الاحتمالي المتصل:

30	.....(ج) - خواص التوقع الرياضي:
30	.....4.1- التباين:
30	.....(ا) - حالة التوزيع الاحتمالي المنفصل:
31	.....(ب) - حالة التوزيع الاحتمالي المتصل:
32	.....5.1-التوزيعات الاحتمالية الشهيرة:
32	.....1.5.1- توزيع ذو الحدين:
33	.....1.1.5.1- خواص توزيع ذو الحدين:
33	.....• التوقع الرياضي:
33	.....• الانحراف المعياري:
33	.....2.5.1- توزيع بواسون:
34	.....1.2.2- خواص توزيع بواسون:
34	.....• التباين:
34	.....• الانحراف المعياري:
34	.....2.2.5.1- تقريب توزيع ذو الحدين من توزيع بواسون:
36	.....3.5.1- التوزيع الطبيعي:
36	.....1.3.5.1- خواص التوزيع الطبيعي:
36	.....• التوقع الرياضي:
36	.....• التباين:
36	.....• الانحراف المعياري:
36	.....2.3.5.1 - التوزيع الطبيعي القياسي:
42	.....3.3.5.1- التقريب الطبيعي لتوزيع ذو الحدين:
44	.....4.5.1- توزيع ستونت (t):
47	.....5.5.1- توزيع كاي مربع ( $\chi^2$ ) khi deux:
48	.....6.5.1- توزيع F (Fischer Snedecor):
49	.....• علاقة بين الـ $\chi^2$ و t و F:

50	6.1- التوزيعات الاحتمالية الثنائية:.....
53	6.1.1- الأمل الرياضي:.....
53	(أ) - حالة المتغيرات العشوائية المنفصلة:.....
53	(ب) - حالة المتغيرات العشوائية المتصلة:.....
53	2.6.1- التباين:.....
53	(أ) - حالة المتغيرات العشوائية المنفصلة:.....
53	(ب) - حالة المتغيرات العشوائية المتصلة:.....
53	3.6.1- التباين المشترك (التغاير):.....
53	(أ) - حالة المتغيرات العشوائية المنفصلة:.....
54	(ب) - حالة المتغيرات العشوائية المتصلة:.....
54	4.6.1- معامل الارتباط:.....
56	تمارين مختارة:.....

## الفصل الثاني

### إختبار الفرضيات

61	مقدمة:.....
62	1.2- أنواع الفرضيات:.....
62	• الفرضية البديلة ذات الذيلين:.....
63	• الفرضية البديلة ذات الذيل الأعلى:.....
63	• الفرضية البديلة ذات الذيل الأدنى:.....
64	2.2- إختبار الأوساط الحسابية:.....
66	1.2.2- إختبار الفرضيات للفرق بين وسطين:.....
68	• إذا كانت $n_1$ و $n_2$ أقل من 30:.....
69	3.2- إختبار التباين:.....
71	4.2- إختبار النسبة بين تباينين:.....
72	5.2- مقارنة بين نسبة نظرية و نسبة مشاهدة:.....

74	.....6.2- إختبار الفرق بين نسبتيين:
75	.....7.2- إختبارات الكاي مربع $\chi^2$ :
	.....1.7.2- مقارنة بين التكرارات المشاهدة مع التكرارات المتوقعة
75	.....(إختبارات المطابقة):
81	.....• تصحيح ياتس Yates:
83	.....2.7.2- إختبارات كاي مربع للإستقلال و التجانس:
86	.....تمارين مختارة:

### الفصل الثالث

#### نظرية العينات

91	.....مقدمة:
91	.....1.3- طرق إختيار العينة:
91	.....1.1.3- طريقة العينة العشوائية البسيطة:
92	.....2.1.3- طريقة العينة الطبقية:
92	.....3.1.3- طريقة العينة العنقودية:
93	.....4.1.3- طريقة العينة المعيارية:
93	.....2.3- توزيعات المعاينة للعينة العشوائية البسيطة:
93	.....1.2.3- توزيعات المعاينة للأوساط الحسابية:
95	.....• حالة عینتین من مجتمعین مستقلین:
96	.....• نظرية النهاية المركزية:
96	.....• حالة مجتمعین مستقلین:
98	.....2.2.3- المعاينة باستعمال توزيع $t$ :
98	.....2.2.3- توزيعات المعاينة للتباينات:
99	.....• حالة $\sigma$ مجهولة:
99	.....3.2.3- توزيعات المعاينة للنسب:
101	.....تمارين مختارة:



## الفصل الرابع التقدير

103	.....مقدمة:
103	.....1.4- معايير جودة التقدير:
103	.....(أ) - عدم التحيز:
104	.....(ب) - التماسك:
104	.....(ج) - الفعالية:
105	.....2.4- التقدير النقطي:
105	.....3.4- التقدير بمجال:
105	.....1.3.4- مجالات الثقة للأوساط الحسابية:
107	.....● حالة $n < 30$ :
108	.....● تقدير الفرق بين وسطين:
109	.....2.3.4- تقدير النسب:
110	.....3.3.4- تقدير الفرق بين نسبتين:
110	.....2.3.4- مجالات الثقة للتباينات:
110	.....(أ) - مجال الثقة لتباين المجتمع:
110	.....(ب) - مجال الثقة للنسبة بين تباينين:
112	.....تمارين مختارة:

## الفصل الخامس نظرية الارتباط

115	.....مقدمة:
115	.....1.5- أنواع العلاقات الارتباطية:
115	.....● علاقات تابعة وارتباطية:
115	.....● علاقة طردية أو عكسية:
116	.....● علاقات مستوية أو منحنية:

116	.....2.5- الارتباط البسيط:
116	.....1.2.5- الارتباط المستقيم و تمثيله:
116	.....1.1.2.5- طريقة المربعات الصغرى:
120	.....2.1.2.5- معامل الارتباط:
120	.....• عبارة أخرى لمعامل الارتباط:
122	.....3.1.2.5- اختبار دلالة معامل الارتباط:
124	.....• حالة اختبار تساوي معاملي ارتباط:
126	.....• اختبار تساوي أكثر من معاملي ارتباط:
127	.....4.1.2.5- معامل التحديد $R^2$ :
128	.....5.1.2.5- دراسة الخطأ المرتكب وتقديره:
130	.....• العلاقة بين معامل الارتباط $r$ ومعامل التحديد $R^2$ :
130	.....6.1.2.5- اختبار معامل التحديد $R^2$ :
130	.....3.5- الارتباط المنحني:
131	.....1.3.5- التمثيل بواسطة معادلة من الدرجة الثانية (معادلة القطع المكافئ):
132	.....2.3.5- التمثيل بواسطة معادلة من الدرجة الثالثة:
133	.....3.3.5- التمثيل بواسطة معادلة لوغاريتمية:
133	.....4.3.5- التمثيل بواسطة معادلة القطع الزائد:
134	.....• المعادلة من الشكل $Y^* = a_0 + \frac{a_1}{X}$ :
134	.....• المعادلة من الشكل: $\frac{X^2}{a^2} - \frac{y^2}{b^2} = 1$ :
136	.....5.3.5- التمثيل بواسطة المعادلة الأسية:
137	.....6.3.5- التمثيل بواسطة المعادلات المثلثية:
138	.....4.5- تطبيقات معادلات التمثيل:
138	.....1.4.5- في مجال الإستطلاع و التنبؤ:
139	.....2.4.5- في مجال حساب المؤشرات الاقتصادية المختلفة:
141	.....5.5- الارتباط المتعدد:

141	.....1.5.5- الخطوات الأساسية لدراسة الارتباط المتعدد:
142	(ا) - إختيار المتغيرات المستقلة و الهامة:.....
144	(ب)- معامل الارتباط المتعدد:.....
145	• حالة متغيرين مستقلين:.....
146	• إختبار دلالة معامل الارتباط المتعدد:.....
146	(ج)- تحديد نوع معادلة التمثيل:.....
147	.....2.5.5- أهم أنواع المعادلات المستخدمة في الارتباط المتعدد:
147	• الارتباط الخطي المتعدد:.....
149	• الارتباط المتعدد ذو الجداء:.....
150	• الارتباط المتعدد الأسّي:.....
151	• الارتباط المتعدد ذو الجداء المختلط:.....
153	.....6.5- الارتباط الجزئي:
154	.....1.6.5- معاملات الارتباط الجزئية:
155	.....2.6.5- تطبيقات الارتباط المتعدد في المجال الاقتصادي:
155	• المرونة الحدية النسبية:.....
155	• المرونة النسبية:.....
155	• المرونة الكلية:.....
155	• المنفعة المتوسطة للمتغير $X$ :.....
156	.....تمارين مختارة:

## الفصل السادس

### السلاسل الزمنية و تحليلها

159	.....مقدمة:
159	.....1.6 - تعريف السلسلة الزمنية:
159	.....2.6-أنواع السلاسل الزمنية:
159	.....3.6 - مؤشرات السلسلة الزمنية:

160	.....(أ) - الزيادة المطلقة:
160	.....(ب) - معدل النمو:
161	.....(ج) - معدل الزيادة النسبية:
161	.....(د) - المتوسط الحسابي لحدود سلسلة زمنية:
161	.....(هـ) - المتوسط التوافقي لحدود سلسلة زمنية:
161	.....(و) - متوسط الزيادة المطلقة:
161	.....(ي) - متوسط معدل النمو:
162	.....(ك) - متوسط معدل الزيادة النسبية:
162	.....4.6- تحليل السلسلة الزمنية:
162	.....1.4.6- تغيرات الاتجاه العام:
163	.....• المتوسطات المتحركة:
165	.....2.4.6- التغيرات الدورية:
165	.....• طريقة الانحرافات النظرية:
165	.....• طريقة المتوسطات المتحركة:
169	.....3.4.6- التغيرات الموسمية:
169	.....4.4.6- التغيرات العشوائية:
170	.....5.6 - التنبؤ:
170	.....1.5.6 - التنبؤ الداخلي:
170	.....2.5.6- التنبؤ الخارجي:
170	.....3.5.6- طرق التنبؤ:
170	.....1.3.5.6 - طريقة التمديد:
171	.....2.3.5.6- طريقة الخطأ المتزايد:
176	.....تمارين مختارة:
179	.....المراجع:
181	.....ملحق الجداول الإحصائية:
189	.....الفهرس:

أنجز طبعه على مطابع

ديوان المطبوعات الجامعية

١، الساحة المركزية - بن عكنون - الجزائر



معتوق أمحمد

# الإحصاء الرياضي والنماذج الإحصائية



ديوان المطبوعات الجامعية